

[illegible]

Db 31 SLPASTLIPVTVICILIFVVGVTGNTMTILLIQYFKDMKTTNLYISSMAVSDVIFLC 90  
 QY 91 LPEFDYLRMRSPWVPGPLCRSLVYVGEQATYLLMTALSVERYLAIICPLRARVLY 150  
 Db 91 LPEFDYLRMRSPWVPGPLCRSLVYVGEQATYLLMTALSVERYLAIICPLRARVLY 150  
 QY 151 TTRRVYALIAVMAVALLSAGPFLVGVGEODPGISVVPGLNGTARJASSPLASSPPLML 210  
 Db 151 TTRRVYALIAVMAVALLSAGPFLVGVGEODPGISVVPGLNGTARJASSPLASSPPLML 210  
 QY 211 SRAAPSPSPSGPETAALALFSRECRSPPAQL--GALRWLMTWTAYEFLPFLCLSTLYG 268  
 Db 186 ---HEDVNTG-----QCKHTGYAISGQHLHIMWSTYTFPCPMCLLFLYG 229  
 QY 269 LIGRELMSRRPLRCPAASGRGHRQTKREVLVAVLAFIICMLPFHVGRITTYNTEDSR 328  
 Db 230 SIGCKLMKSGNDIGCALARESHROTIVKLVVAVLAFIICMLPFYHIGRMLFAQVDYD 289  
 QY 329 MMVFSQYFNIVALQFLYLSASINPILYNLSKRYRAAFKLLARKSRPRGPHRSR 384  
 Db 290 TMLISQNFMAVSLCYLSASINPVVYNLMSRKYRAAKRLFLHQ-RPKPAHRGQ 344

RESULT 2  
 093412 PRELIMINARY; PRT; 374 AA.

AC 093412; 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)  
 DE Orphan G protein-coupled receptor.  
 OS Spherozooids nephelus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Spherozooids.  
 NC NCB1\_TaxID=39110;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20092336; PubMed=10628755;  
 RA Palyna O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,  
 Rao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,  
 Patchett A.A., Howard A.D., Smith R.G.;  
 RT "Ligand activation domain of human orphan growth hormone (GH)  
 RL Mol. Endocrinol. 14:160-169(2000).  
 DR EMBL: AF082209; AAC33472.1;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR InterPro: IPR000524; HTH\_GntR.  
 DR Pfam: PF00001; 7tm\_1.1.  
 DR PRINTS: PR00237; GPCR\_RHODPSN.  
 DR PRINTS: PR01565; NEURONMEDINUR.  
 DR PRINTS: PR01566; NEURONMEDINUR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR PROSITE: PS00043; HTH\_GNTR\_FAMILY; UNKNOWN\_1.  
 DR KW Receptor.  
 SQ SEQUENCE 374 AA; 42324 MW; 2CF9304FD004C7A16 CRC64;

Query Match 39.2%; Score 845; DB 13; Length 374;  
 Best Local Similarity 45.5%; Pred. No. 3.6e-63;  
 Matches 187; Conservative 64; Mismatches 102; Indels 58; Gaps 12;

QY 5 WNGS-DGEGAREPPWALPDCDRCSPFLGALVPTAVCLCFVVGSGNVVTVMIL 63  
 Db 13 WEGSHNGTAGLELP-----LNYISIPLAIVTACTVLPVGVGAVMTILV 61  
 QY 64 GRVDMRTTNLYIGSMVSDLLILGLPFDLYLRMRSPWVPGPLCRSLVYVGEQATY 123  
 Db 62 SRYRMRRTTNLYLCSMAVSDLFIFVCMPLDLYRMKRYRPMRGDLYCKLFQVVSCTY 121  
 QY 124 ATLLMTALSVERYLAIICPLRARVLYTTRRVYALIAVMAVALLSAGPFLVGVGEODP 183

Db 122 STILICTALSVERYLAIICPLRAKALVTKRRVRLIILMTVSLISAGPVVMNGEKDS 181  
 QY 184 GISVVPGLNGTARJASSPL-ASSPPLMLSRAPPSPPSGPETAALALFSRECRPS--PA 240  
 Db 182 -IMFNP-----SSDLNESSWPL-----EADTRECNRQYAV 212  
 QY 241 QLGLRWLMTWTAYEFLPFLCLSTLYGLIGRELMSRRPLR-GPAASGRGHRQTKRV 299  
 Db 213 ESGLEAMVWLSSVFFPMVPCVLTVGLIGRLMLRRETTINGRVAVRDKSNRQTIKM 272  
 QY 300 LVVVLAFFICMLPFHVGRITTYNTEDS--RMVFSQYFNIVALQFLYLSASINPILYN 356  
 Db 273 LVVVLAFFICMLPFHVGRITTYNTEDS--RMVFSQYFNIVALQFLYLSASINPILYN 332  
 QY 357 LISKRYRAAFKLLARKSRPRGPHRSRDTAGEVAGTGGDTGVYETTSAN 407  
 Db 333 TMSWKYRGAVARLFGVSDSPQ--RGR-TASTVMD-----GMTSTVTS 373

RESULT 3  
 093414 PRELIMINARY; PRT; 559 AA.

AC 093414; 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)  
 DE Orphan G protein-coupled receptor.  
 OS Spherozooids nephelus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Spherozooids.  
 NC NCB1\_TaxID=39110;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20092336; PubMed=10628755;  
 RA Palyna O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,  
 Rao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,  
 Patchett A.A., Howard A.D., Smith R.G.;  
 RT "Ligand activation domain of human orphan growth hormone (GH)  
 RL Mol. Endocrinol. 14:160-169(2000).  
 DR EMBL: AF082211; AAC33474.1;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1.1.  
 DR PRINTS: PR00237; GPCR\_RHODPSN.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECP\_F1\_2; 1.  
 DR KW Receptor.  
 SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAF0CD5F6 CRC64;

Query Match 27.9%; Score 602; DB 13; Length 559;  
 Best Local Similarity 29.5%; Pred. No. 1.6e-42;  
 Matches 155; Conservative 65; Mismatches 125; Indels 180; Gaps 12;

QY 22 LPPCDRCR---SPFLGALVPTAVCLCFVVGSGNVVTVMILGRVDMRTTNLYIG 78  
 Db 20 LHKCSNQCWHEBPV-FGMIVCVTIIVPLMLFGLIGNIITLVVWMLRPMYRSSTVYLS 78  
 QY 79 SMAVSDLLILGLPFDLYLRMRSPWVPGPLCRSLVYVGEQATYLLMTALSVERYL 138  
 Db 79 SLAVSDILILLPLVDLYKLMRPMRPMFGEIFCKSTYFSECTFCSILMTITFLSLRYL 138  
 QY 139 AICRPLRARVLYTTRRVYALIAVMAVALLSAGPFLVGE-----QDPTIS----- 186  
 Db 139 AVCMPTAKVTVTRRTTRITIGICWGAISAPVVMVGVGEVQSDQGISGWRSGAW 198  
 QY 187 -----VVPGL----- 191  
 Db 199 TGKEGEGFIIIGRERENDKGLKDBQLEBMMWKKXEMNCGDKNGVTKGKGDKSLIV 258  
 QY 192 -NGTARJASSPLASSPPLMLSRAPPSPPSGP-----TAAALALFSRECRPS--PAQLGA 244

[illegible]

```

Db      61  MICAATYLLIFVYVAGVNGGLCVILIRKAKMRFTYNYLPSLAVSDLLVLLGLPELYE 120
QY      98  LMRSPFWFGPCLLCRLSYVEGGTYATLLMTALSVERYLALICPLRPAVLVYRRVRVA 157
Db     121  MMHNYPFLIGVGCGCYFTLLPEMYCLASLVNTALSVERYAAVNHPIQASMSMTBAHYR 160
QY     156  LIALVMAVALLSAGPFLFVVEGDDPGISVYVGLNGTARIASSPLASSPMLSRADPPS 217
Db     181  VLGAWVGAMLCSPENLTHGIRO--LHV-----20707
QY     218  PSCGETAEAAALFESRCRSPAPQALRVMLWVTA--YFPLPCLSTLTXGLIGRELM 275
Db     208  PGRGVVPDPAICMLVLP-----RALYNNVQVTTALLFCPLMAISVLYLLIGRLR 259
QY     276  SSRRLP-----RGPAA-----GREGRHQRTRVLVVLVLAFLICWLPFHVGRIT 320
Db     260  REFLLMQEAKGGRSAARSRYTCRLQHQHRRGRQVTKMLFVLVVVFGICMAFPHADRW 319
QY     321  Y--INTESRMMYSQVFNIVALQFLPLSINFIYLNLSKKYRPAAPFL---LARK 374
Db     320  MSVVSQMTDGHIALFOHVHVISGIFPYLGSANPVLISLMSRFRFPQALCLAGACR 379
QY     375  SRPR--GFHRSRDTAGEVAAGDTG 395
Db     380  LRPRHSHSLSRMTTGSTLCDVG 402

```

ID	Q91276	PRELIMINARY;	PRT;	395 AA.
AC	Q91276;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Neuromedin U receptor type 2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RA	Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,			
RA	Gustafson E.L.;			
RT	"Characterization of murine neuromedin U R2 receptor";			
RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AY057384; AL26695.1; -			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm.1.1.			
DR	PRINTS; PRO1565; NEUROMEDINUR.			
DR	PRINTS; PRO1567; NEUROMEDINUR.			
DR	PRINTS; PRO1570; NPFFRECEPTOR.			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1.1; UNKNOWN_1.			
DR	PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.			
KW	Receptor.			
SC	SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;			
Query Match	22.1%;	Score 475.5;	DB 11;	Length 395;
Best Local Similarity	30.6%;	Pred. No. 5e-12;		
Matches 115;	Conservative 79;	Mismatches 101;	Indels 81;	Gaps 11;
QY	30 CSPPEPLGALVYTVAVCLFLPVGSGVGNVYMLGRYDMKMTTNLYGSAVSDLI-L 88			
Db	29 CGPGRSDLSLPSVSVVALILFVVGIVIGMLVCLVLAHQTLTPNNYVYFSLAVSDLIYL 88			
QY	89 LGLEFDLYRLMRSRPWFGPLLCRLSLVYEGECYVATLHMTALSVERYLAICRPLRARV 148			
Db	89 LGMPLEVEYELMHNPRLFEGVPGCYFKTALPFTVGFASILSTVYISIEKYVALIVDFRAKL 148			
QY	149 LVYRRRRALIVLMAVALISAGFFLVLGV--EODPGISVVGELNGTARTIASSPLASSP 206			
Db	149 ESTRRRLRLTLISLWSPVSLPNTSIHGKFKQDFPNGSGVPG-----SATCTVTK 200			

```

Oy 207 PLMLSRAPPPSPSGPETAAMALFSRECRPSDAQALRMLMTTAAFF-LPFLCLST 265
Db 201 PMWV-----YNFIIQATSFYIIPMTLISV 226
Oy 266 LYLIGREL-----WSSRRPLRGPAASGRGRHGRKRVLLVLAFTICWLP 313
Db 227 LYLMLGLRLKRDSELEADKVTYVIRPSR-----KSVTKMLFVLVLAFTICWLP 275
Oy 314 FHVGRHIVYNTEDSRMWFVSQYFN--IVALQLFYLSASINPILYMLISKRYAAAFKLL 370
Db 276 FHVDR-LFSPFVDEWETESLAAYVNLHVVSQVFYLLSSAVNPITLNLSSRFR-AAFRNV 333
Oy 371 LA-----RKSRRPG 379
Db 334 VSPSCKCHPHRPOG 349

RESULT 6
Oy 096AM5 PRELIMINARY; PRT; 415 AA.
ID 096AM5;
AC 01-DEC-2001 (Tremblrel, 19, Created)
DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Neuromedin U receptor 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC016938; AAI16938.1;
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 1
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1570; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor
SQ SEQUENCE 415 AA; 47770 MW; 30BFDD706436AB9 CRC64;

Query Match 22.0%; Score 475; DB 4; Length 415;
Best Local Similarity 30.8%; Pred. No. 5.8e-32;
Matches 115; Conservative 78; Mismatches 98; Indels 82; Gaps 12;

Oy 30 CSPFPLGALVAVTAVCLFVVGSGVNVVTMLIGRYRDMRTTNLYLGSMVSDLL-L 88
Db 37 CGPRRSHEFLVSVVYVPIFVVGIGVNLVCLVILQHQMKTPTNYLFLSLAVSDLVLL 96
Oy 89 LGLPFDLYRLMRSPVWVGPLLGRSLVYEGCTYATLLHMTALSVERYLAICPLRAV 148
Db 97 LGMFLVLEVEMRNRYFLGPGVCYFKTALFETVCPASISITVSEYVAILHPRFRL 156
Oy 149 LVTRRRVALAVLMAVALISAGPFLVGV--QDPGIVVPGVNGTARIASSPLASSP 206
Db 157 QSTRRRRLRIIGIIVGFSVLFSLPNTSIHGKHFPPNSLVPG-----SATTVYK 208
Oy 207 PLMLSRAPPPSPSGPETAAMALFSRECRPSDAQALRMLMTTAAFF-LPFLCLST 265
Db 209 PMWV-----YNFIIQATSFYIIPMTLISV 224
Oy 266 LYLIGRELWSSR-----RPLRGPAASGRGRHGRKRVLLVLAFTICWLP 313
Db 235 LYLMLALMLKKKSLADGEGNANIGPCR-----KSVTKMLFVLVLAFTICWLP 283
Oy 314 FHVGRHIVYNTEDSRMWFVSQYFNVAL--QLFYLSASINPILYMLISKRYAAAFKLL 369
Db 284 FHIDRLFSPFVSEWSEST--AAVFNLHVVSQVFYLLSSAVNPITLNLSSRFR-AAFON 340

```

```

Oy 370 LLARKSRPRGFHR 382
Db 341 VIS-----SPHK 347

RESULT 7
Oy 09GZ04 PRELIMINARY; PRT; 415 AA.
ID 09GZ04
AC 09GZ04; 09NR06;
DT 01-MAR-2001 (Tremblrel, 16, Created)
DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Neuromedin U receptor 2 (Neuromedin U receptor-type 2) (G
DE protein-coupled receptor TGR-1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artyushyn R., Bonini J.A., Borowsky B.,
RA Dole J.L.W., Zhou S., Koutanova E.V., Nigorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vayse P.J., Branchek T.A., Gerald C., Forray C.,
RA Adam N.;
RT "Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System";
RL J. Biol. Chem. 275:32452-32459 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RL PubMed=11010960;
RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
RA Gustafson E.L., Monma F.J. Jr., Hedrick J.A.;
RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in
RT the Central Nervous System";
RL J. Biol. Chem. 275:39482-39486 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RL Pang L., Wang S., Laz T., Hedrick J.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 4-415 FROM N.A.
RL MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Peng S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA Stral J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.,
RA "Identification of receptors for neuromedin U and its role in
RT feeding";
RL Nature 406:70-74 (2000).
RN [5]
RP SEQUENCE OF 4-415 FROM N.A.
RL PubMed=10887190;
RA Hosoya M., Moriya T., Kawabata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neuromedin U Receptor";
RL J. Biol. Chem. 275:29528-29532 (2000).
RL EMBL: AF272363; AAG24794.1;
DR EMBL: AF292402; AAG03064.1;
DR EMBL: AF242874; AAF82755.1;
DR EMBL: AB041228; BAB13721.1;
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 1
DR PRINTS; PRO0237; GPCR_Rhodopsin.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1570; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_2; 1.

```



DR PROSITE, PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW RECEPTOR.  
 SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B31390F CRC64;

Query Match 22.0%; Score 474; DB 4; Length 415;  
 Best Local Similarity 30.8%; Pred. No. 7e-32;  
 Matches 115; Conservative 78; Mismatches 98; Indels 82; Gaps 12;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILIGRYDMRTTNLYLGSMAVSDLLI-L 88  
 DB 37 CGFRSRHFLPVSVAVALIFLVGMGNLVCWIVAHQTLKPTNNYLLFSLAVSDLLVL 96  
 QY 89 IGLPDLIRLMSRPWVFGPLLCLSLYVEGCTVATLHMTALSVERYLAICRPLRARV 148  
 DB 97 LGMPLIEVEMHNYPLFEGVGCYFETALFETVCFAISLITVSERYVALIHPFRKL 156  
 QY 149 LVTRRRVALLIIVMAVALLSAGPFLVGV--QDPGISVVPGLNGTARIASSPLASSP 206  
 DB 157 QSTRRRALRIIGIVMGSVFLSPLENTSIHGIRKFFNGSLVPG-----SATCTYIK 208  
 QY 207 PLMLSRAPPSPSPGPTAAALFSRECRPSPAQLGALRVMLWTTAYFF-LPFLCLSI 265  
 DB 209 PMWV-----YNLIQATSFLLPMTLVISV 234  
 QY 266 LYGLIGRELMSR-----PRLRGPASGRGRGHQT-KRVLLVLAFTICWLP 313  
 DB 225 LYVLMALRLKDKSLKDEADGNANIQRPCK-----KSVNKMFLVILVFAICWAP 283  
 QY 314 PHVGIITYNTED-SRMWYFSQYFNIVAL--QLFYLSASINPLIYNLISKRYRAARLLA-----RKS 375  
 DB 284 FHIDRLFFSFEVSEWSESL--AAVENLVHVVSQVFFYLSSAVNPITLYLSRRFQ-AAFQV 340  
 QY 370 LLARKSRPRGHR 382  
 DB 341 VIS-----SFHK 347

## RESULT 8

Q9JTB1 PRELIMINARY; PRT; 395 AA.

AC Q9JTB1; 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Neuromedin U receptor 2.  
 GN NMU2R.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=20351041; PubMed=10894543;  
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,  
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,  
 RA Strait J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McEneaney K.K.,  
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austen C.P.,  
 RA Caskey T., van der Ploeg L.H.T., Liu Q.,  
 RT "Identification of receptors for neuromedin U and its role in  
 feeding.";  
 RL Nature 406:70-74 (2000).  
 DR EMBL; AF242875; AAF82756.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PRO00237; GPCRHOPOPSN.  
 DR PRINTS; PRO1565; NEUROMEDINUR.  
 DR PRINTS; PRO1567; NEUROMEDINUR.  
 DR PRINTS; PRO1570; NEUROMEDINUR.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;

Query Match 21.9%; Score 472.5; DB 11; Length 395;  
 Best Local Similarity 30.5%; Pred. No. 8.9e-32;  
 Matches 111; Conservative 81; Mismatches 115; Indels 57; Gaps 9;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILIGRYDMRTTNLYLGSMAVSDLLI-L 88  
 DB 29 CGFRSRHFLPVSVAVALIFLVGMGNLVCWIVAHQTLKPTNNYLLFSLAVSDLLVL 88  
 QY 89 IGLPDLIRLMSRPWVFGPLLCLSLYVEGCTVATLHMTALSVERYLAICRPLRARV 148  
 DB 97 LGMPLIEVEMHNYPLFEGVGCYFETALFETVCFAISLITVSERYVALIHPFRKL 148  
 QY 149 LVTRRRVALLIIVMAVALLSAGPFLVGV--QDPGISVVPGLNGTARIASSPLASSP 206  
 DB 149 ESTRRRALRIISLVVSFVSFLPNTSIHGIRKFFNGSLVPG-----SATCTYIK 200  
 QY 207 PLMLSRAPPSPSPGPTAAALFSRECRPSPAQLGALRVMLWTTAYFF-LPFLCLSI 265  
 DB 201 PMWV-----YNLIQATSFLLPMTLVISV 226  
 QY 266 LYGLIGRELMSR-----PRLRGPASGRGRGHQT-KRVLLVLAFTICWLP 313  
 DB 227 LYVLMALRLKDKSLKDEADGNANIQRPCK-----KSVNKMFLVILVFAICWAP 283  
 QY 325 ED--SRMWYFSQYFNIVAL--QLFYLSASINPLIYNLISKRYRAARLLA-----RKS 375  
 DB 287 EHWTESLAAVNPLIHVVSQVFFYLSSAVNPITLYLSRRFQ-AAFQV 340  
 QY 376 RPRG 379  
 DB 346 RPOG 349

## RESULT 9

Q9ESQ4 PRELIMINARY; PRT; 395 AA.

AC Q9ESQ4; 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE G protein-coupled receptor TGR-1.  
 GN TGR-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20449029; PubMed=10887190;  
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,  
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,  
 RA Fujino M.,  
 RT "Identification and Functional Characterization of a Novel Subtype of  
 Neuromedin U Receptor.";  
 RL U. Biol. Chem. 275:29528-29532 (2000).  
 DR EMBL; AB041229; BAB13722.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PRO00237; GPCRHOPOPSN.  
 DR PRINTS; PRO1565; NEUROMEDINUR.  
 DR PRINTS; PRO1567; NEUROMEDINUR.  
 DR PRINTS; PRO1570; NEUROMEDINUR.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 395 AA; 44756 MW; EABD6D36B3C355DA CRC64;

Query Match 21.8%; Score 470.5; DB 11; Length 395;  
 Best Local Similarity 31.2%; Pred. No. 1.3e-31;  
 Matches 107; Conservative 75; Mismatches 112; Indels 49; Gaps 7;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILIGRYDMRTTNLYLGSMAVSDLLI-L 88

```

Db 29 CGPRSDSLSTPVSVAVALIFLVGMGNLLVCMVYVHQTLPKPTNYLFLSLAVSDLLVL 88
QY 89 LGLPEDLRLMRSPWVFGPLICRLSLVYEGCTATLHMTALSVERYLAI CRPLRAV 148
Db 89 LGMPLEIEMHNYFLPGVPCYFKTALFEIVCFASISLVTSEYVAI VHPFRAL 148
QY 149 LVTRRRVALIAVMAVALISAGPLFLVGV--EODPGISVVPGLNGTARIASSPLASSP 206
Db 149 ESTRRRLRLISLWVSFVSFLPTNSIHGKQHPFGSSVPG-----SATCTVTK 200
QY 207 PLMLSRAPPPSPSGPETAALFSECRPPAQCALRVMMLVTAYFF--LPFLCLSI 265
Db 201 PMWV-----NLIIGATSLFLYLLPMTLISV 226
QY 266 LYLIGRELMSRRPLRGPAASGRGRHQT--KRVLLVVLVLAFLICMLPHVGRILYINT 324
Db 227 LYLMLGLRLKRDESLANKVAVNIHPRSKSVTKMLFVLVLPALCWTDPHVDLRFESFV 286
QY 325 ED--SRMWFSGYFNIVALQFLTSLASINPIIYNLISKKRYAA 365
Db 287 EEMTESLAAVFNLIVHVSQVGFYLLSSAVNPDIYNLISRRFRAA 329

```

## RESULT 10

055040 PRELIMINARY; PRT; 405 AA.

```

ID 055040;
AC 055040;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN Orphan G protein-coupled receptor.
GN GPR66.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feigheimer S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D.,
RA Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RL EMBL: AF044602; AAC02681.1; -.
DR MGD; MGI:11341898; GPCR.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
KW SEQUENCE 405 AA; 45609 MW; F1BA493DBB81F34 CRC64;

```

Query Match 21.2%; Score 456.5; DB 11; Length 405;  
 Best Local Similarity 32.2%; Pred. No. 2e-30;  
 Matches 133; Conservative 62; Mismatches 141; Indels 77; Gaps 14;

```

QY 19 WPAIPLP---CDE---RRCSPFGLALVPYAVCLCFVGVSGNVVVMIGRYRDMR 70
Db 10 WPYQPEDLNLDEALRLKYLIGPOKQKQFPCVITYLLIFVGLGGLCTYILRKTKR 69
QY 71 TTTNYLGSMAVSDLLIL--GLPFDLYRLMRSPWVFGPLICRLSLVYEGCTATLHM 129
Db 70 TPTNYFLSLAVSDMLVLIVGLPLEIYEQQVYFPQLGASACVFRILLLETVCASLVNV 129
QY 130 TALSVERYLAI CRPLRAVLTTRRRVALIAVMAVALISAGPLFLVGVGEODPGISVVP 189
Db 130 TALSVERYVAVVRPLQAKSVMTTRAHVRMVGAIWVLAFLSLPNTSIHGLSQ---LTVP 185
QY 190 GLNGTARIASSPLASSPLMLSRAPPPSPSGPETAALFSECRPPAQCALRVM 249
Db 186 -----CRGPVPD-----SAIGS---LVGMDVYKL--VVL 210

```

```

QY 250 WTTAYFFFLCLSLIYLIGRELMSRRPL-----RGPASG-----REGRHR 294
Db 211 TTALEFCLPMVTIVSYLLIIGLRLRRERMLQVEVGRKTAATQETSHRRIQLQDGRR 270
QY 295 QTRVLLVVLVLAFLICMLPHVGRILYI-----NTEDSRMWFSGYFNIVALQFLYLSA 349
Db 271 QVTKMLFALVVVFGICWAPFHADRIMSLVYGHSTBELHLAY--QCVHISGIFLYGSA 328
QY 350 INPIIYNLISKKRYAAAFKL-----LARKSPRGFRS--RDTAGEVAGDGT 395
Db 329 ANPVLVLSMSTRFRFTFLQALGLTQCCHRRQPYHGSHNIRLTGTTGTLCDVG 381

```

## RESULT 11

017239 PRELIMINARY; PRT; 418 AA.

```

ID 017239;
AC 017239;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 47.7 kDa protein.
GN K10B4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RA "Genome sequence of the nematode C. elegans: a platform for
RA investigating biology. The C. elegans Sequencing Consortium. ";
RT Science 282:2012-2018(1998).
RL [2]
RN RN
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RA Sammons L., Wohlmann P., Antonion B.;
RT "The sequence of C. elegans cosmid K10B4. ";
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission. ";
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF025463; AAB71009.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Hypothetical protein.
KW SEQUENCE 418 AA; 47735 MW; CE0416539CA3BE27 CRC64;

```

Query Match 20.1%; Score 434; DB 5; Length 418;  
 Best Local Similarity 30.1%; Pred. No. 1.7e-28;  
 Matches 109; Conservative 77; Mismatches 138; Indels 38; Gaps 10;

```

QY 29 RCPFPGLALVPYAVCLCFVGVSGNVVVMIGRYRDMRTTNYLGSMAVSDLL-I 87
Db 24 RCQ--SAGIIVPIVITYGTFLLGLFNGICTCIYIANKSMHNPTNYLFLSLAVSDIIL 81
QY 88 LGLPFDLYR-LMRSPWVFGPLICRLSLVYEGCTATLHMTALSVERYLAI CRPLRA 146
Db 82 ILGLPMEFYQSLDYSYRSEEGICKARAFLEFTSYASIMTICFSFERMLAICHPURS 141
QY 147 RVLVTRRRVALIAVMAVALISAGPLFLVGVGEODP---GISVVPGLNGTARIASSPLA 203
Db 142 KISTIMRANVLLIILATWISVCLPIAFIVQINKLPLPEDAKXQPMWNKVSFAVGVIN 201

```

QY 204 SSPLMSTRAPPSPGSPETAEAAALFSREC-----PSPQAOLCALRYMLWTTAYEFL 258  
 Db 202 -----NRIFFPV-----TDGIFVHTETFCAMNQSRPDQKM-----IIFAFVTFVFI 244  
 QY 259 PFLCLSTLYGLIGRELMSSRRPLRGP--AAGRGHGHQTRKVLVVLAFICMLPHV 316  
 Db 245 PALAIVMVAIAVQLSSSEIDLKDDKVKKRNRKNSKRTVLKMLLSVITFPCMLPHFI 304  
 QY 317 GRITTYNTEDSRMVFQSYFNIVALQLEFYL-----ASINPILYNLISKRYRAAPFL 369  
 Db 305 QRLLSVYTTWSETTITSPVQFLMIVYISGFCYSSAANPILYNLISQRYSAFCRT 364  
 QY 370 LL 371  
 Db 365 IL 366

RESULT 12  
 Q9VFNA PRELIMINARY; PRT; 428 AA.  
 AC Q9VFNA; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE CG9918 protein.  
 GN CG9918.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Bokoyva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kailush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazolo M., Peltman G.S., Pan S., Pollard D.R., Puri V., Reese M.G.,  
 RA Rainer K., Remington K., Saunders R.D.C., Schebler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).

CC -- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL: AB03703; AAF55016.1; --  
 DR FLYbase: FBgn0038201; CG9918.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm 1; 1.  
 DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
 DR PROSITE: PRO1565; NEUROKININOR.  
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR\_F1\_1; 1.  
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.  
 SQ SEQUENCE 428 AA; 47198 MW; 9DD4A1E9304686D7 CRC64.

Query Match 19.9%; Score 430; DB 5; Length 428;  
 Best Local Similarity 27.7%; Pred. No. 3.7e-28;  
 Matches 119; Conservative 61; Mismatches 119; Indels 130; Gaps 10;

QY 23 PCDERRCSFPFLGALVPTAVCLCFVVGSGNVTVWMLIGRDMRTTNLYLSMAV 82  
 Db 12 PPRD-----PLAIVPTVYSLIFITGVGNISTCIVIKKRSMTATNYLFSLAI 64  
 QY 83 SDLLLL-GLPPDLYRLMSRPWFGLPLCRSLVYEGCTYATLHMTALSVRYLAIC 141  
 Db 65 SDFLLLSGVDPQEVSYMSKYVYFGEYICIGRLAETSANATVLTITAFYERYIAIC 124  
 QY 142 RPLRAVLVTRRRVALLAVLMAVALLAGPFLVLVGQDQGISVPGANGTARLASSP 201  
 Db 125 HPLFGQMSKSRARITVLMVIMAVITAIPOAOFGLHYSGVE----- 169  
 QY 202 LASSPPLWLSRAPPSPPSGPETAEEAALFSRECRPSPQALALRYML---WVTAYPF 257  
 Db 170 -----QCIIVYIVHSGQLSTIFTF 190  
 QY 258 L-PFLCLSTLYGLIGRELMSSRRPLRGPASGRER----- 291  
 Db 191 LAPSIITLVLLIGVHLRYSTL-VEGPASVARRQQLSVPSDTILRYGSGTAMSFNG 249  
 QY 292 -----GHRTKRVLLVVLAFICMLPHVGHIIYI- 322  
 Db 250 GGSAGTAGLGMGSGAQLSSVRGLNHTGRVRLMLVAVAVCFPLCWAPFAOIRIATY 309  
 QY 323 -----NTEDSRMVFQSYFNIVALQLEFYLSSASINPILYNLISKRYRAAPFLIARKSR 376  
 Db 310 APARGAKLRDQHEFVYV-VMTYVSGVLYISTCINPLINMHSFR-BAFRAVLFGKV 367  
 QY 377 PRGFRSRD 385  
 Db 368 SKGSLNSRN 376

RESULT 13  
 Q9VFNA PRELIMINARY; PRT; 595 AA.  
 AC Q9VFNA; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE CG8795 protein.  
 GN CG8795.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Zhang Q., Pfeiffer B.D.,



Db 348 LGGTRCCHRHQPR 361

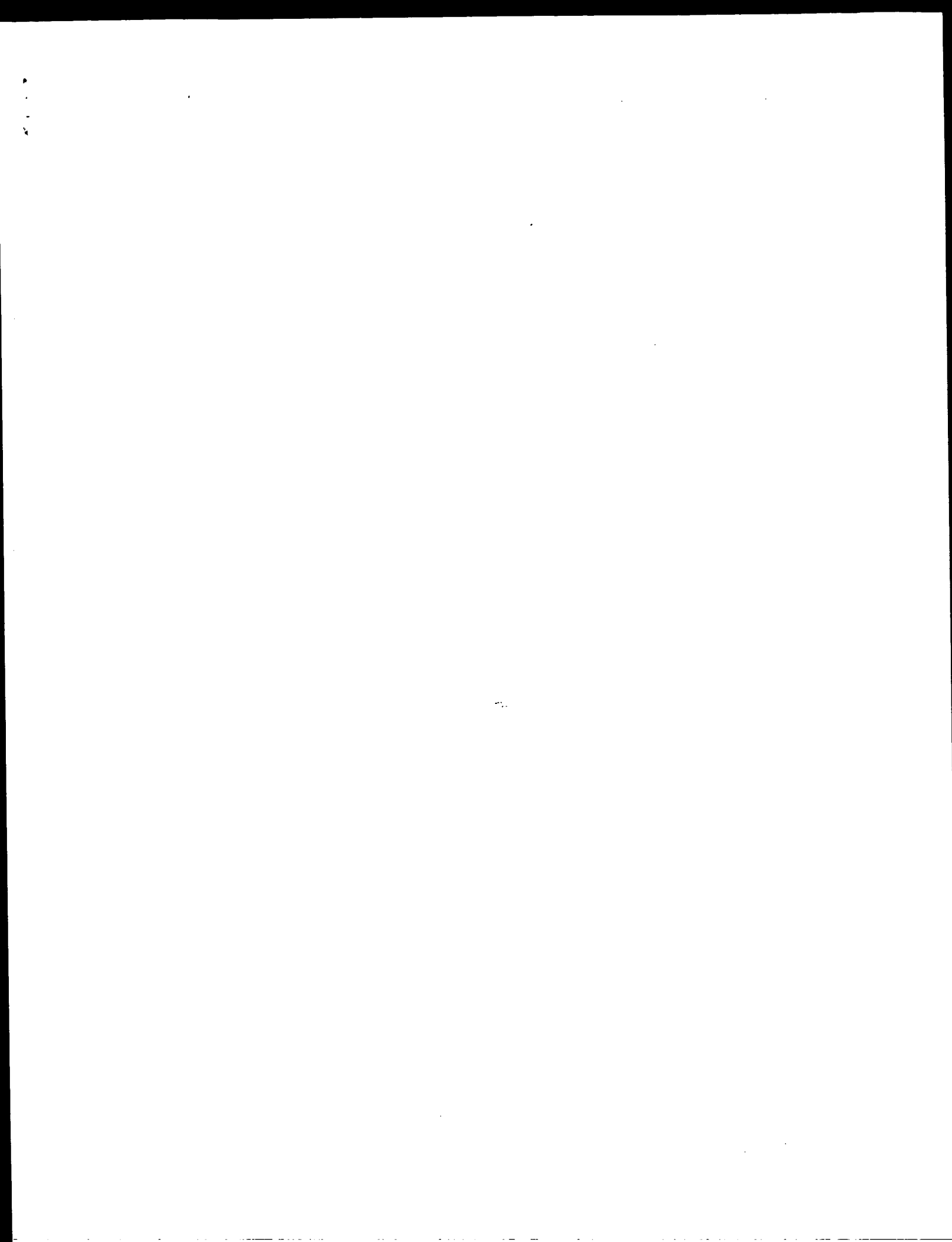
## RESULT 15

08TBH6 PRELIMINARY; PRT; 410 AA.  
 AC 08TBH6;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Neurotensin receptor 2.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Strausberg R;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC022501; AAH22501.1; -  
 KW Receptor.  
 SQ SEQUENCE 410 AA; 45459 MW; DC3ADA4B6D58D7BD CRC64;

Query Match 18.8%; Score 405; DB 4; Length 410;  
 Best Local Similarity 30.9%; Pred. No. 4.5e-26;  
 Matches 128; Conservative 61; Mismatches 135; Indels 90; Gaps 16;

QY 17 PPMFALP---CDERCSPPFLGALVPTAVCLTFVVGSGNVVTWLI-----GR 65  
 DB 8 PPRSSNPGLSLDARLGVDRLWAKVLFYALYALIMALGAAGNALSVHVLKARAGRAGR 67  
 QY 66 YRDMRTTNLYLGSMAVSDLLIL-GLPDLXR-LMRSRPWFEGPILCRSLVYEGGCTY 123  
 DB 68 LRH-----HYLSLALAGLLLVGPVELYSFVWFHPVFGDLGCRGYFVHELCA 120  
 QY 124 ATLLHMTALSVRYLAICRPLRARVLTTRRRVALLVMAVALISGPFLELVG----- 179  
 DB 121 ATVLSTAGLSAERCLVACQPLRAKSLTTPRTTLMVALSWAASLGLMPMAVIMGQKH 180  
 QY 180 -----EQDGIISVPGINGTARIASS-----PLASSPPL-----WLS 211  
 DB 181 ETADGEPBPASRYCTVL--VSRVALQVFIQVNVLVSVFLPLALTAFUNGVTSHLALCS 238  
 QY 212 RAPPPSPGSPETAAALPFSRECRPSPAOLGALRVMLWVTYVFLPLCLSLYGLIG 271  
 DB 239 QVFSSTTP-GSSTPSRLLELSE-----GLISFIWKKT-----FIG--G 276  
 QY 272 RELWSSRRPLRGPAASGRGRGHROTCKVLLVVLAFICWLPHVGRITINTEDSR--- 328  
 DB 277 QVSLVRHKDVR-----RIRSLQRSYQVLAIVVMYVICMLPYHARLMICYVPDDAMTD 330  
 QY 329 -MMYFSQYFNIALQLFYLSASINPLILYNLISKRYRAAAFKLLARKSRPRGPH 381  
 DB 331 PLYNFYHYFVMTNLFYVSSAVTPLLNVAVSSSFR-----RLFLNAVSSSLGSH 380

Search completed: March 16, 2003, 15:05:03  
 Job time : 37 secs



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:10 ; Search time 13 Seconds  
(without alignments)  
1314.481 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156  
Sequence: 1 MGSFNGSDGPEGAREPWP.....DTGGDTVGTETSTANVTMG 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2149	99.7	412	1	MTLR_HUMAN
2	858.5	39.8	366	1	GHSR_PIG
3	857.5	39.8	366	1	GHSR_HUMAN
4	855.5	39.7	364	1	GHSR_PAT
5	614	28.5	257	1	GHSR_MOUSE
6	498	23.1	424	1	NTR1_PAT
7	495.5	23.0	424	1	NTR1_MOUSE
8	491	22.8	418	1	NTR1_HUMAN
9	407	18.9	410	1	NTR2_HUMAN
10	404	18.7	417	1	NTR2_MOUSE
11	401	18.6	416	1	NTR2_PAT
12	369.5	17.1	362	1	SSRS_MOUSE
13	358	16.6	362	1	SSRS_PAT
14	356.5	16.5	418	1	SSRS_HUMAN
15	349.5	16.2	453	1	GP39_HUMAN
16	343	15.9	364	1	SSRS_HUMAN
17	342.5	15.9	380	1	OPRX_CAVPO
18	341.5	15.8	370	1	OPRX_CAVPO
19	338	15.7	370	1	OPRX_PAT
20	334	15.5	380	1	OPRX_PAT
21	334	15.5	388	1	OPRX_HUMAN
22	333.5	15.5	367	1	SSR4_HUMAN
23	333.5	15.5	367	1	OPRX_MOUSE
24	333.5	15.5	370	1	OPRX_PAT
25	333.5	15.5	428	1	SSR3_PIG
26	332.5	15.4	372	1	SSR3_PAT
27	331.5	15.4	395	1	OPRX_HUMAN
28	331	15.4	428	1	SSR3_MOUSE
29	330.5	15.3	370	1	OPRX_HUMAN
30	329.5	15.3	370	1	OPRX_MOUSE
31	328	15.2	380	1	OPRX_MOUSE
32	326.5	15.1	371	1	GALS_MOUSE
33	326.5	15.1	384	1	SSR4_PAT

34	323.5	15.0	370	1	GALT_PAT	088626	rattus norv
35	321	14.9	384	1	SSR4_MOUSE	P49660	mus musc
36	320.5	14.9	372	1	OPRX_PAT	P33533	rattus norv
37	320.5	14.9	398	1	TRPR_HUMAN	P34981	homo sapien
38	320	14.8	519	1	TLR2_DROME	P06755	drosophila
39	319.5	14.8	368	1	GALT_HUMAN	060755	homo sapien
40	319	14.8	372	1	OPRX_MOUSE	P32300	mus musc
41	319	14.8	453	1	CCR_XENLA	P70031	xenopus lae
42	318.5	14.8	387	1	GALS_HUMAN	043603	homo sapien
43	318.5	14.8	398	1	TRPR_BOVIN	046639	bos taurus
44	318	14.7	353	1	APJ_XENLA	P79660	xenopus lae
45	318	14.7	380	1	APJ_HUMAN	P35414	homo sapien

## ALIGNMENTS

RESULT 1  
ID MTLR\_HUMAN STANDARD: PRT: 412 AA.  
AC Q43153.

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE Motilin receptor (G protein-coupled receptor GPR38).

GN MTLR1 OR MTLR OR GPR38.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID:9606;

[1] SEQUENCE FROM N.A. (ISOFORM A).

RA MEDLINE=9810578; PubMed=9441746;

RA McKeen K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,

RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;

RT "Cloning and characterization of two human G protein-coupled receptor

RT genes (GPR38 and GPR39) related to the growth hormone secretagogue

RT and neurotensin receptors.";

RL Genomics 46:426-434(1997).

[2] SEQUENCE FROM N.A. (ISOFORMS A AND B).

RA MEDLINE=99316084; PubMed=10381885;

RA Feighner S.D., Tan C.P., McKee K.K., Palyna O.C., Hreniuk D.L.,

RA Pong S.-S., Austin C.P., Figueroa P., MacNeil D., Cascieri M.A.,

RA Nargund R., Bakhshi R., Abramowitz M., Stocco R., Kargman S.,

RA O'Neill G., van der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,

RT "Receptor for motilin identified in the human gastrointestinal

RT system.";

RL Science 284:2184-2188(1999).

[3] SEQUENCE FROM N.A. (ISOFORM A).

RA Wall M.

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

[4] FUNCTION.

RA MEDLINE=21219832; PubMed=11322507;

RA Smith R.G., Leonard R., Bailey A.R.T., Palyna O.C., Feighner S.D.,

RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;

RT "Growth hormone secretagogue receptor family members and ligands.";

RL Endocrine 14:9-14(2001).

CC -1- FUNCTION: Receptor for motilin.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are

CC produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE

CC MARROW.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way





FT DISULFID 116 198 BY SIMILARITY.  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 266 289 AVVFAFLLCMLPFHVRGYFSKS -> GGSQCLLESLPG  
 FT VARSPLIC 290 366 PLSCLFSSP (IN ISOFORM 1B).  
 SQ SEQUENCE 366 AA, 41194 MW, 2C850B35F61B7C1C CRC64;  
 Query Match 39.8%; Score 858.5; DB 1; Length 366;  
 Best Local Similarity 44.9%; Pred. No. 1.6e-48;  
 Matches 188; Conservative 55; Mismatches 105; Indels 71; Gaps 10;  
 QY 5 WNGSDGDEGA-----REPMPALPCCD---BRCSPPPLGALVVTVAVCLCFVGVGSGN 56  
 DB 2 WATPEEPEGPNLTLPDLCMDAPPENDSLVEELLPLPPTPLGAVTATCVAFVVGIAGN 61  
 QY 57 VVTVMICGRVDRMTTNTLYLGMAVSDLLTLGLPFDLYRMRPMVREGPLLCRLSLV 116  
 DB 62 LITMLVVSFRERMTTNTLYLSMAVSDLLTLGLPFDLYRMRPMVREGPLLCRLSLV 121  
 QY 117 VEEGCTATLHMTALSVRYIAICRPLRARVLTERRRVALIAVMAVALLSAGPEFL 176  
 DB 122 VSESCYATVLTITALSVERVYFAICRPLRARVLTERRRVALIAVMAVALLSAGPEFL 181  
 QY 177 VGEVDPGISVPLGNGTARINASSPLASSPMLSRAPPSPGPTAAALFSECR 236  
 DB 182 VGEVHD-----NGT-----DPRD-----TNECR 199  
 QY 237 PS-PAOLGALRMLVMTAVFPLPCLSLYLGLIGRELMSRRPLRGAASG---RER 291  
 DB 200 ATEFAVRSGLLTVVMVSSVFFFLPVFCLTVLSLGRKM---RRKRGAAVSSLRDQ 256  
 QY 222 GHRQKRVLLVVAFLICMLPPHVRITV---INTEDSHMYFSQYFNIALQLFYLSA 348  
 DB 257 NHRQVVKMLAVVFAFLLCMLPFHVRGYFSKSLEPGSVIAQISQYCNLVSVFLFYLSA 316  
 QY 349 SINPLIYMLISKYVAAAFKLLARSRPGRHRSRDTAGEVADGCDGVGTETSAN 407  
 DB 317 AINPLIYMLISKYVAAAFKLLARSRPGRHRSRDTAGEVADGCDGVGTETSAN 365  
 RESULT 3  
 GHSR HUMAN STANDARD; PRT; 366 AA.  
 AC Q92847; Q92848; Q96R07;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing  
 peptide receptor) (GHRP) (ghrelin receptor).  
 GN GHSR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
 RC TISSUE=pituitary;  
 RX MEDLINE=9633798; PubMed=8688086;  
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,  
 RA Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,  
 RA Palyna O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,  
 RA McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevics M.,  
 RA Pechelet A.A., Nargund R., Srinatchsingh D.J.S., Dean D.C., Meillich D.G.,  
 RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;  
 RA "A receptor in pituitary and hypothalamus that functions in growth  
 hormone release.";  
 RT hormone release.";  
 RL hormone release.";  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
 RX MEDLINE=21255649; PubMed=11356716;  
 RA Peterseim S., Rasch A.C., Penhorn M., Bell F.U., Schulte H.M.;

RT "Genomic structure and transcriptional regulation of the human growth  
 hormone secretagogue receptor.";  
 RT Endocrinology 142:2649-2659(2001).  
 RL [3]  
 RN FUNCTION.  
 RP MEDLINE=21219832; PubMed=11322507;  
 RA Smith R.G., Leonard R., Bailey A.R.T., Palyna O.C., Feighner S.D.,  
 RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;  
 RA "Growth hormone secretagogue receptor family members and ligands.";  
 RT Endocrine 14:9-14(2001).  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE=20067959; PubMed=10604470;  
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;  
 RA "Ghrelin is a growth-hormone-releasing acylated peptide from  
 stomach.";  
 RT Nature 402:656-660(1999).  
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.  
 CC Stimulates growth hormone secretion. Binds also other growth  
 hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)  
 CC as well as non-peptide, low molecular weight secretagogues (e.g.  
 CC L-692,429, MK-0677, adenosine).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND  
 CC SECRETAGOGUES.  
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U60179; AAC50653.1; -;  
 DR EMBL; U60181; AAC50654.1; -;  
 DR EMBL; AF369786; AK71539.1; -;  
 DR EMBL; AF369786; AK71540.1; -;  
 DR Genew; HGNC:4267; GHSR.  
 DR MIM; 601898; -;  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS00237; G-PROTEIN RECP\_F1\_1;  
 DR PROSITE; PS00262; G-PROTEIN RECP\_F2\_1;  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Alternative splicing.  
 FT DOMAIN 1 40  
 FT TRANSMEM 41 66  
 FT DOMAIN 67 72  
 FT TRANSMEM 73 96  
 FT DOMAIN 97 117  
 FT TRANSMEM 118 139  
 FT DOMAIN 140 162  
 FT TRANSMEM 163 183  
 FT DOMAIN 184 211  
 FT TRANSMEM 212 235  
 FT DOMAIN 236 263  
 FT TRANSMEM 264 285  
 FT DOMAIN 286 302  
 FT TRANSMEM 303 326  
 FT DOMAIN 327 366  
 FT DISULFID 116 198  
 FT CARBOHYD 13 13  
 FT CARBOHYD 27 27  
 FT VARSPLIC 266 289  
 FT VARSPLIC 290 366  
 SQ SEQUENCE 366 AA, 41328 MW, D1B62710DA9DC0C6 CRC64;

Query Match 39.8%; Score 857.5; DB 1; Length 366;  
 Best Local Similarity 44.5%; Pred. No. 1.9e-48;  
 Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

5 WNGSGPGGA-----REPWPALPCD---ERRCSPPFLGALVPTAVCLCFVVGSGN 56  
 2 WNAAT--PSEEPBNVTLTDLMDASFGNDSPDLPLFPAPLAGVATCVAFVVGISG 61  
 57 VVTWMLIGRYDRMTTNTLYLGSMVSDLLILGLPFDIRLMSRPVWFGPLCRSLY 116  
 62 LITMLVSRFRRLRTTNTLYLSMAFSDLLIFLCMPDLVRLMQRPVWFGDLCLCFQ 121  
 117 VGEGCTYATLLMTALSVRYALICRPLARVLYRRRRAIALVMAVALSAGPFL 176  
 122 VSESCYATVLTITLALSVRYALICRPLAKVVTGRVLYIVMAVAFSAGTIFVL 181  
 177 VGEVODPGISVVGNGTARIASSPLASPPMLSRAPPSPGPETAAMALFSRRC 236  
 182 VVEVHE-----NGT-----DP--W-----DINECR 199  
 237 PS--PAQLGALRYMLWTTAYFFLPCLSLIYGLIGRELMSSRRLRGPASGRERGR 294  
 200 PFEAFVRSGLITVMVWVSSIFFLPVFCITVLSLIGRKLRRRGDAVAGSLRDQNHK 259  
 295 QTRKVLVAVVLAFLICMLPFRVGRITVINTEDS---RMVYFSQYFNIVLQLFYLSASIN 351  
 260 QTVKMLAVVFAFLICMLPFRVGRILFSKFEFGSLIAQISQYCNLVSVLYLSA 319  
 352 PLYNLSKKYRAAFLLARSRPRGFRSRDTAGVAGDGTGVTYETSAN 407  
 320 PLYNLSKKYRAVAVRLLGFEPFSQKSLTLDSESR-----AWTESIN 365

RESULT 4  
 GHSR RAT STANDARD; PRT; 364 AA.

AC 008725;  
 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).  
 GN GHSR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pituitary;  
 RX MEDLINE=9724655; PubMed=9092793;  
 RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P., Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;  
 RA "Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors";  
 RT Mol. Endocrinol. 11:415-423 (1997).  
 RN [2]  
 RP SEQUENCE OF 1-240 FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Pituitary;  
 RX MEDLINE=98100386; PubMed=9437732;  
 RA Yokoe R., Sato M., Matsubara S., Ohye H., Niimi M., Murao K., Takahara J.;  
 RA "Molecular cloning and gene expression of growth hormone-releasing peptide receptor in rat tissues";  
 RT Peptides 19:15-20 (1998).  
 RN [3]  
 RP FUNCTION.  
 RX MEDLINE=20067959; PubMed=10604470;  
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;  
 RA "Ghrelin is a growth-hormone-releasing acylated peptide from stomach.";  
 RT Nature 402:656-660 (1999).  
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.

CC Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine) (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-stb.ch/announce/> or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).  
 CC  
 CC EMBL; U94321; AAC53156.1; -.  
 CC EMBL; AB001962; BA01777.1; ALT\_INIT.  
 CC InterPro; IPR000276; GPCR\_Rhodopsn.  
 CC Pfam; PF00001; 7tm\_1; 1.  
 CC PRINTS; PR00237; GPCR\_Rhodopsn.  
 CC PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 CC PROSITE; PS0262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein.  
 CC DOMAIN 1  
 CC TRANSMEM 41 66  
 CC FT 67 72  
 CC DOMAIN 67 72  
 CC TRANSMEM 73 96  
 CC FT 97 117  
 CC DOMAIN 118 139  
 CC TRANSMEM 140 162  
 CC FT 163 183  
 CC DOMAIN 184 211  
 CC TRANSMEM 212 235  
 CC FT 236 263  
 CC TRANSMEM 264 285  
 CC FT 286 302  
 CC TRANSMEM 303 326  
 CC FT 327 364  
 CC DOMAIN 364 370  
 CC DISULFID 115 197  
 CC FT 13 13  
 CC CARBOHYD 26 26  
 CC FT 26 26  
 CC SEQUENCE 364 AA; 40963 MW; DCBF55BEO61BEE9 CRC64;

Query Match 39.7%; Score 855.5; DB 1; Length 364;  
 Best Local Similarity 48.0%; Pred. No. 2.5e-48;  
 Matches 184; Conservative 52; Mismatches 82; Indels 65; Gaps 11;

5 WNGSDPBGARREP-----PWPALPCD---ERRCSPPFLGALVPTAVCLCFVVGSG 55  
 2 WNAAT--PSEEPBNVTLTDLMDASFGNDSPDLPLFPAPLAGVATCVAFVVGISG 59  
 56 NVVTWMLIGRYDRMTTNTLYLGSMVSDLLILGLPFDIRLMSRPVWFGPLCRSL 115  
 60 NLTMLVSRFRRLRTTNTLYLSMAFSDLLIFLCMPDLVRLMQRPVWFGDLCLCFQ 119  
 116 VGEGCTYATLLMTALSVRYALICRPLARVLYRRRRAIALVMAVALSAGPFL 175  
 120 VSESCYATVLTITLALSVRYALICRPLAKVVTGRVLYIVMAVAFSAGTIFV 179  
 176 VGEVODPGISVVGNGTARIASSPLASPPMLSRAPPSPGPETAAMALFSRRC 235  
 180 LVGEVHE-----NGT-----DP--W-----DINECR 197  
 236 RPS--PAQLGALRYMLWTTAYFFLPCLSLIYGLIGRELMSSRRLRGPASG--RE 290  
 198 RATEFAVRSGLITVMVWVSSIFFLPVFCITVLSLIGRKLW--RR--RDAAVAGSLRD 253  
 291 RGHQTRKVLVAVVLAFLICMLPFRVGRITVINTEDS---RMVYFSQYFNIVLQLFYLS 347  
 254 QNHQTVKMLAVVFAFLICMLPFRVGRILFSKFEFGSLIAQISQYCNLVSVLYLS 313  
 348 ASINPIYNLISKYRAAFLL 370

Db 314 AAINPLNINMSKRVAVFKLL 336

RESULT 5  
GHSR\_MOUSE  
ID GHSR\_MOUSE STANDARD; PRT; 257 AA.  
AC Q99P50; Q91282; (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).  
GN GHR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 1-183 FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;  
RA Kacsoh B.;  
RT "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by rapid amplification of cDNA ends (RACE)."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 73-257 FROM N.A.  
RC STRAIN=129S3/SvImJ;  
RA Peng X., Frohman L.A., Kineman R.D.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine) (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AY056474; AAL1336.1; -  
CC EMBL: AF332997; AAC6114.1; -  
CC InterPro: IPR000276; GPCR\_Rhodopsn.  
CC Pfam: PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 35  
FT TRANSSEM 36 58  
FT DOMAIN 59 77  
FT TRANSSEM 78 100  
FT DOMAIN 101 125  
FT TRANSSEM 126 148  
FT DOMAIN 149 160  
FT TRANSSEM 161 183  
FT DOMAIN 184 211  
FT TRANSSEM 212 234  
FT DOMAIN 235 257  
FT DISULFID 115 197  
FT CARBOHYD 13 13  
FT CARBOHYD 26 26  
FT CARBOHYD 187 187  
FT NON\_TER 257  
SQ SEQUENCE 257 AA; 28743 MW; D6FA21EAB0E30791 CRC64;  
Query Match 28.5%; Score 614; DB 1; Length 257;  
Best Local Similarity 45.1%; Pred. No. 5.7e-33;

Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

QY 5 WNGSDGPEGAREP-----PWPALPCD---ERRCSPPGLGVVTVAVCLCFVGVG 55  
Db 2 WNAAT--PSEEPENVTLDLDWDASPGNDSDELPLPEPAPLAVATCAVAFVGGIS 59  
QY 56 NVVTMLIGRYDRMTTNLYIGSMANVSDLLILGLPFDLYRLMSPVFPVPLCRSL 115  
Db 60 NLTLMLVSRRELRTTNLYLSSMAFSDLLIFLCMPDLVRLMYRPNFDDLCKLFQ 119  
QY 116 YVGGCTYATLTHNTALSVERYLAICRPLARVLTTRRRVALLAVMAVALLAGPFLF 175  
Db 120 FVSECTYATVLTITLALSVERYFAICFPLRAKVYTKGRVLIIVIAVAFCSAGPIFV 179  
QY 176 LVGEOPDGISVPLNGLTARLASSPLASSPPLMSTRAPPSPGPETAFAALFSPRC 235  
Db 180 LVGVEHE-----NGT-----DPRD-----TWEC 197  
QY 236 RPS--PAQGLRVMLWTTAVFPLPFLCLSLVGLIGRELMSRRPLRGPAASG---RE 290  
Db 198 RATEFAVRSGLLTVMWVSVFFLPVPCLTIVLSLRKLM--RR--RGDAVGSRLRD 253  
QY 291 RGRH 294  
Db 254 QNHK 257

## RESULT 6

ID NTR1\_RAT STANDARD; PRT; 424 AA.

AC P20789;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Neurotensin receptor type 1 (NTR-1) (High-affinity leucobastine-insensitive neurotensin receptor) (NTRH).  
GN NTRSL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90297956; PubMed=1694443;  
RA Tanaka K., Masu M., Nakanishi S.;  
RT "Structure and functional expression of the cloned rat neurotensin receptor."  
RL Neuron 4:847-854(1990).  
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC HIGHEST TO TACHYKININS RECEPTORS.  
CC PIR; JH0164; JH0164.  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR InterPro: IPR003985; NTL\_Rec.  
DR InterPro: IPR003984; NTL\_Rec.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PRINTS; PR01479; NEUROTENSINR.  
DR PRINTS; PR01480; NEUROTENSINR.  
DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 64  
FT TRANSSEM 65 87  
FT DOMAIN 88 96  
FT TRANSSEM 97 121  
FT DOMAIN 122 143  
FT TRANSSEM 144 165

FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 189 210 4 (POTENTIAL).  
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 236 260 5 (POTENTIAL).  
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 309 330 6 (POTENTIAL).  
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 349 372 7 (POTENTIAL).  
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 142 225 BY SIMILARITY.  
 FT LIPID 388 388 PALMITATE (POTENTIAL).  
 SQ SEQUENCE 424 AA, 47054 MW, A9C2F7A9B9BCD3 CRC64;

Query Match 23.1%; Score 498; DB 1; Length 424;  
 Best Local Similarity 32.9%; Pred. No. 2.5e-25;  
 Matches 125; Conservative 58; Mismatches 111; Indels 86; Gaps 11;

QY 39 VPTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LGLGFPD 94  
 DB 65 VLTAVLYALFVVGTVGNSVTAFLLARKKSLQSLQSTVHYHGLSLSDLLILLAMPVE 124  
 QY 95 LVR-LMRSRPWFPGPLCLSLYVGGCTYATLLHMTALSVERYLAIICRPLARAVLTRR 153  
 DB 125 LYNFTVWHHPMAFGDAGCGRYFLRDCTATALNVASLSYERLAIICHPFKATLMSRS 184  
 QY 154 RVRALIAVMAVALLSAGPFLFVVGEDPGISVVPGLNGTARIASSPLASSPPLMLSRA 213  
 DB 185 RTKKFISATLWALSALLAIPMLFTWGLQNRSGDGTGHPG---GLVCTPIVDVT----- 231  
 QY 214 PPPSPSPGPTAALALFSRECRSPAOGLARMLMTTAYFFL-PFLCLSLIYGLIGR 272  
 DB 232 -----ATVKKVIVQNTFMSFLPMLVLSINTVTIAN 262  
 QY 273 ELMSRRPLRGPAASGR-----ERGHROTGR-----VLTAVLYALFI 309  
 DB 263 KLTVM---VHQAAEQGVCTVGTNGLEHSTFNMITEGRVQALRHGLVLRVAVIAFVV 319  
 QY 310 CWLPFHVGRITTYINTEDSR---MMYFSQYENIVALOLFYLASINPILYNLSIKKYRAA 365  
 DB 320 CWLPFHVGRITTYINTEDSR---MMYFSQYENIVALOLFYLASINPILYNLSIKKYRAA 365  
 QY 366 AFKLLA-----KRSRP 377  
 DB 380 FLSTLACLCPGWRHRRKKRP 399

RESULT 7  
 NTRL MOUSE STANDARD; PRT; 424 AA.  
 AC 088319;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurotensin receptor type 1 (NT-R-1).  
 GN NTSR1 OR NTSR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Snider J., Sano H., Ohta M.;  
 RT "Neurotensin receptor type 1,"  
 RT "Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases."  
 RL - FUNCTION: RECEPTOR FOR THE TRIDECAPPTIDE NEUROTENSIN. IT IS  
 CC ASSOCIATED WITH G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
 CC CALCIUM SECOND MESSENGER SYSTEM.  
 CC - SUBCELLULAR LOCATION: Integral membrane protein.  
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

HIGHEST TO TACHYKININS RECEPTORS.  
 CC EMBL; AB017827; BAB3013.1; -  
 CC MGD; MGI:97386; Ntsr.  
 CC Interpro; IPR000276; GPCR\_Rhodpsn.  
 CC Interpro; IPR003985; NT1\_rec.  
 CC Interpro; IPR003984; NT1\_rec.  
 CC Pfam; PF00001; 7tm.1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN.  
 DR PRINTS; PR01479; NEUROTENSINR.  
 DR PRINTS; PR01480; NEUROTENSINR.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 65 87 1 (POTENTIAL).  
 FT TRANSMEM 88 96 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 97 121 2 (POTENTIAL).  
 FT TRANSMEM 122 143 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 144 165 3 (POTENTIAL).  
 FT TRANSMEM 166 188 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 189 210 4 (POTENTIAL).  
 FT TRANSMEM 211 235 5 (POTENTIAL).  
 FT TRANSMEM 236 260 6 (POTENTIAL).  
 FT TRANSMEM 261 308 7 (POTENTIAL).  
 FT TRANSMEM 309 330 8 (POTENTIAL).  
 FT TRANSMEM 331 348 9 (POTENTIAL).  
 FT TRANSMEM 349 372 10 (POTENTIAL).  
 FT TRANSMEM 373 424 11 (POTENTIAL).  
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 141 224 BY SIMILARITY.  
 FT LIPID 388 388 PALMITATE (POTENTIAL).  
 SQ SEQUENCE 424 AA, 47216 MW, 859A73171A48711 CRC64;

Query Match 23.0%; Score 495.5; DB 1; Length 424;  
 Best Local Similarity 33.6%; Pred. No. 3.6e-25;  
 Matches 127; Conservative 56; Mismatches 114; Indels 81; Gaps 12;

QY 39 VPTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LGLGFPD 94  
 DB 64 VLTAVLYALFVVGTVGNSVTAFLLARKKSLQSLQSTVHYHGLSLSDLLILLAMPVE 123  
 QY 95 LVR-LMRSRPWFPGPLCLSLYVGGCTYATLLHMTALSVERYLAIICRPLARAVLTRR 153  
 DB 124 LYNFTVWHHPMAFGDAGCGRYFLRDCTATALNVASLSYERLAIICHPFKATLMSRS 183  
 QY 154 RVRALIAVMAVALLSAGPFLFVVGEDPGISVVPGLNGTARIASSPLASSPPLMLSRA 213  
 DB 184 RTKKFISATLWALSALLAIPMLFTWGLQNRSGDGTGHPG---GLVCTPIVDVT----- 231  
 QY 214 PPPSPSPGPTAALALFSRECRSPAOGLARMLMTTAYFFL-PFLCLSLIYGLIGR 272  
 DB 214 ADGQHGCG-----LVCTPT-VDTATVAVVIVQNTFMSFLPMLIISINTVTIAN 261  
 QY 273 EL---WSSRRPLRGPAASGR-----ERGHROTGR-----VLTAVLYALFI 311  
 DB 262 KLTVMHQAABEQRGVCTVGTNGLEHSTFNMITEGRVQALRHGLVLRVAVIAFVV 321  
 QY 312 LPHVGRITTYINTEDSR---MMYFSQYENIVALOLFYLASINPILYNLSIKKYRAAF 367  
 DB 322 LPHVGRITTYINTEDSR---MMYFSQYENIVALOLFYLASINPILYNLSIKKYRAAF 367

OY 368 KLLA-----RKSRP 377  
DB 382 STIACLCGWRRRRRKKRP 399

## RESULT 8

NTRI\_HUMAN STANDARD; PRT; 418 AA.  
AC P30989; Q9H4T5; Q9H4H1;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucocastine-  
sensitive neurotensin receptor) (NTRH).  
OS NTSR1 OR NTR.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93154505; PubMed=8381365;  
RA Vita N., Laurent P., Lefort S., Chalon P., Dumont X., Kaghad M.,  
RA Gully D., Le Fur G., Ferrara F., Caput D.;  
RT "Cloning and expression of a complementary DNA encoding a high  
RT affinity human neurotensin receptor";  
RL FEBS Lett. 317:139-142(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill M.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clegg S.,  
RA Clegg S., Cooley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A.G., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
RA Ellington D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Garner P.,  
RA Graffham D.V., Griffiths J.A., Fraser A., French L., Garner P.,  
RA Hammon S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levesalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McComachie L.J., McKay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,  
RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20";  
RL Nature 414:865-871(2001).  
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS  
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
CC CALCIUM SECOND MESSENGER SYSTEM.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC HIGHEST TO TACHYKININS RECEPTORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X70070; CAA49675.1; --  
DR EMBL; AL357033; CAC14923.1; --  
DR EMBL; AL035669; CAC12747.1; --  
DR PIR; S29506; S29506.  
DR HSP; P02699; 1F88.  
DR Genew; HGNC:8039; NTSR1.  
DR MIM; 162651; --  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR InterPro; IPR003985; NT1\_Rec.  
DR InterPro; IPR003984; NT\_Rec.  
DR Pfam; PF00001; 7tm\_1; 2.  
DR PRINTS; PR01479; NEUROTENSINR.  
DR PRINTS; PR01480; NEUROTENSINR.  
DR PROSITE; PS00237; G\_PROTEIN\_RECIP\_F1\_1; 1.  
DR PROSITE; PS50262; G\_PROTEIN\_RECIP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 63  
FT TRANSMEM 64 86  
FT DOMAIN 87 95  
FT TRANSMEM 96 120  
FT DOMAIN 121 142  
FT TRANSMEM 143 164  
FT DOMAIN 165 187  
FT TRANSMEM 188 209  
FT DOMAIN 210 234  
FT TRANSMEM 235 259  
FT DOMAIN 260 303  
FT TRANSMEM 304 325  
FT DOMAIN 326 343  
FT TRANSMEM 344 367  
FT DOMAIN 368 418  
FT CARBOHYD 4 4  
FT CARBOHYD 37 37  
FT CARBOHYD 41 41  
FT DISULFID 141 224  
FT LIPID 383 383  
FT CONFLICT 200 200  
SQ SEQUENCE 418 AA; 46288 MW; BBBDECB2B6E390 CRC64;  
  
Query Match 22.8%; Score 491; DB 1; Length 418;  
Best Local Similarity 32.8%; Pred. No. 6; 9e-25;  
Matches 125; Conservative 62; Mismatches 118; Indels 76; Gaps 13;  
  
OY 39 VPTAVCLCFVGVGSGNVTVNLIGR---YRDRRTTNLYLSMAVSDIL-IILGLPFD 94  
DB 64 VLTAVYALFVGVGVGNVTATFTLARKKSLOSQSTVHVLGSLASDLTLTLMPVE 123  
  
OY 95 LVR-LWRSRPVFGPLLCLSLVVGECYATLHMTALSEVRYLATCRPLARVLVTR 153  
DB 124 LVNFIWVHHFWAGDAGCRGVYFLRDACTATLNLVNSLSEVRYLATIHPKATIMRS 183  
  
OY 154 RVVALAVLMAVALLSAGPFLVGVHODGIVSPGLNGTARIASSPLASSPPLWSRA 213  
DB 184 RTKFFISAIWLSALTLVPLFTMG--EON----- 211  
  
OY 214 PRSPSPSGPETAALALFSECRSPAPQAGALRVLMTT-AVFLFLCLSLVGLIGR 272  
DB 212 -----RSADGQAGAGVCTPT-IHTATVKVIVQIVNTFMGFIIPMVVISTNTIAN 261  
  
OY 273 ELWSSRRPL--RGPAS-----GREGHRTKRVLLVNVLAFLICWLPFHV 316  
DB 262 KLTVMVROAAEQGVCTVGGESHSTFMAIRGVRVQALRHGVRLRAVVAVFVVCWLPFHV 321  
  
OY 317 GRITVINTEDSR---MMY-FSQYFNVALQFLYLASINPLVLSKKYR-----AA 365  
DB 322 RRLMFCYISDEQWTPPLPYFYHYFVWTVNALFYVSSITINILVNSANPRHIFLATLAC 381  
  
OY 366 AFKLLARKSRPGFHSRSDT 386  
DB 382 LCPVWRKKRKP-AFSRKADS 401

```

Matches 129; Conservative 60; Mismatches 135; Indels 90; Gaps 16;
QY 17 PEPNDALP---CBERKSPPLGALVPAVAVCCLELVGVGVGVVVVVMI-----GR 65
Db 8 PPRSSNPGLSLDARIGVDTRIMAKVLFATLVALGAGNALSVHYVLRARAGR 67
QY 66 YRDMRTTNLYLGSMVAVDLIL-L-GLPEDLVR-LWRSRPWGPRLICRLSLVYGGCTY 123
Db 68 LRH-----HYSLALAGLLLVGVPELVSEFWPHYMWVGDDCGRGYYFVHELCAV 120
QY 124 ATLLMTALSVERYLAIQRLPRARVIVTRRRVALIAVMVALSLSGPFLVGV----- 179
Db 121 ATVLSTVAGLSAERCLAVCGQLRARSLTPRRTRKWLVALSWAASLGLALPMVAVIGQKHEI 180
QY 180 -----EODPGISVYPGINGTARIAS-----PLASPPV-----MTS 211
Db 181 ETADGEEPPASRVCYTV--VSRTALQVFIQVNVNVLSFVLEPLATLFLNGVTSHLLALCS 238
QY 212 RAPDPSPGSGEYVTEAALFSPREGCPSPDAGLRVWLWTTAYFFLPFLCLSLVGLIG 271
Db 239 QVPSSTIP-GSSTPSRELLSEE-----GLSLFVWKKT-----FIQG--G 276
QY 272 RELWSSRRPLRGPAASRGRRGHOTKRVLLVVLAFICMLPEPHVGHIIYINEDSR--- 328
Db 277 QVSLVRHKDYR-----RISLGRSVQLRAIVMVCWLPYHARIMCYVPDDAMTD 330
QY 329 -MMTFSQYFNVLVALQFLVASINPLIYNLSKKYRAARLLARSRGRGFH 381
Db 331 PLNYFYHYFYVWNTLFFVSSAVPLLYNAVSSFR---KLPLEAVSLGHEH 380

RESULT 10
ID NTNR2_MOUSE STANDARD; PRT; 417 AA.
AC P70310;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurotensin receptor type 2 (NT-R-2) (low-affinity leucobastine-
DE sensitive neurotensin receptor) (NTR2).
GN NTR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxId=10090;
RN NCBI_TaxId=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=96388216; Pubmed=8795617;
RA Mazella J, Botto J.-M., Guillemare E., Coppola T., Sarret P.,
RA Vincent J.-P.;
RT "Structure, functional expression, and cerebral localization of the
RT leucobastine-sensitive neurotensin/neuromedin N receptor from mouse
RT brain.";
RL J. Neurosci. 16:5613-5620(1996).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDCAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM,
CC HIPPOCAMPUS, PIRIFORM CORTEX AND NEOCORTEX OF ADULT BRAIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN.
CC EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-
CC OLD BRAIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation /
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)

```

or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch).

CC -----  
 CC EMBL; U51908; AAB17285.1; -  
 DR MGD; MGI:108018; Ntsr2.  
 DR InterPro; IPR000276; GPCR Rhodopsin.  
 DR InterPro; IPR003986; NT2\_rec.  
 DR InterPro; IPR003984; NT2\_rec.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PRO1479; NEUR0TENSINR.  
 DR PRINTS; PRO1481; NEUR0TENSINR.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KM G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.  
 FT DOMAIN 1  
 FT TRANSMEM 33  
 FT DOMAIN 56  
 FT TRANSMEM 64  
 FT DOMAIN 87  
 FT TRANSMEM 88  
 FT TRANSMEM 109  
 FT TRANSMEM 110  
 FT TRANSMEM 132  
 FT TRANSMEM 155  
 FT TRANSMEM 176  
 FT TRANSMEM 218  
 FT TRANSMEM 239  
 FT TRANSMEM 298  
 FT TRANSMEM 320  
 FT TRANSMEM 339  
 FT TRANSMEM 360  
 FT TRANSMEM 417  
 FT DISULFID 108  
 FT LIPID 378  
 FT SEQUENCE 417 AA; 46537 MW; EBRDDB6507223DD CRC64;

Query Match 18.7%; Score 404; DB 1; Length 417;  
 Best Local Similarity 30.6%; Pred. No. 2.6e-19;  
 Matches 126; Conservative 58; Mismatches 146; Indels 82; Gaps 14;

QY 19 WPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTWMLIGRYDMRT- 71  
 Db 7 WPPRPSGAGSLERLGLVDTRLMAKVFETALYSILFAGTGNALSHVVLKARTGRPG 66  
 QY 72 TTNYLIGSMAYSDLLIL-GLPFDLYR-LMRSRPWFGLCLSLYVGGCTYATLLH 129  
 Db 67 RLRYHVLSTLALLLLLSVMEIYFWSHYPWFGLGCRGYFVVELCAIYAVLSV 126  
 QY 130 TALSVETLALCRPLARVLTTRRVALIAVLMVALLSAGPFLVGVGDDPGISVVP 189  
 Db 127 ASLSAERCLAVCOPLRARLLTPRRCTRLSLVWVASLGALPMAYIMQKHE----- 179  
 QY 190 GLNGTARIASSPLASSPPLMSRAP-PPSPGSPETAALALFSRECRSPALQ---GAL 245  
 Db 180 -----MERADGEPPAS-----RVCTVIVSRASSRSSTQVKRAGLL 215  
 QY 246 RVLMLVTTA-----YFPLPFLC-----LSIL--YGLIGREIMSSRPL 281  
 Db 216 RSPLEMLTALINGITVNLVALYISQVPSASQVNSIPSLLELSEBGLLGFTWRTLSL 275  
 QY 282 RGPASAGREGRHQCK-----RVLLVVVLATTCMLPHVGRITTYNTD---SRMY 331  
 Db 276 GVQASLVNHHKASQTRLSQHSQAQVLAIVAVVICMLPYHARLWVCYLPDGGWTDLYD 335  
 QY 332 FSOYFNIALQLFYLSASINPILYVLSKYYRAAPFKLLARKSPRGHRS 383  
 Db 336 FYHIFVWVNTLTFVSSAVTVLVNAVSSFR---KLFLESLSLCGEORS 383

RESULT 11  
 NTR2 RAT  
 ID NTR2 RAT STANDARD; PRT; 416 AA.  
 AC 063384;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurotensin receptor type 2 (NT-R-2) (High-affinity leucocabastine-

DE sensitive neurotensin receptor).

GN NTR2 OR NTR2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Hypothalamus;

RX MEDLINE=96228041; PubMed=8647296;

RA Chalon P., Vite N., Kaghad M., Guillemont M., Bonin J.,  
 RA Delpech B., Le Fur G., Ferrara P., Caput D.;  
 RT "Molecular cloning of a leucocabastine-sensitive neurotensin binding  
 RT site.";

RT FEBS Lett. 386:91-94(1996).

-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUR0TENSIN. IT IS  
 ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
 CALCIUM SECOND MESSENGER SYSTEM.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER  
 LEVELS SEEN IN THE HEART AND INTESTINE.

-1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND  
 EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD  
 BRAIN).

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 HIGHEST TO TACHYKININS RECEPTORS.

-----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.sib-sib.ch/announce/>  
 or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC EMBL; X97121; CA65787.1; -  
 CC InterPro; IPR000276; GPCR Rhodopsin.  
 CC InterPro; IPR003986; NT2\_rec.  
 CC InterPro; IPR003984; NT2\_rec.  
 CC Pfam; PF00001; 7tm\_1; 1.

DR PRINTS; PRO1479; NEUR0TENSINR.  
 DR PRINTS; PRO1481; NEUR0TENSINR.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.  
 FT DOMAIN 1  
 FT TRANSMEM 33  
 FT DOMAIN 56  
 FT TRANSMEM 64  
 FT DOMAIN 87  
 FT TRANSMEM 88  
 FT TRANSMEM 109  
 FT TRANSMEM 110  
 FT TRANSMEM 132  
 FT TRANSMEM 155  
 FT TRANSMEM 176  
 FT TRANSMEM 217  
 FT TRANSMEM 237  
 FT TRANSMEM 297  
 FT TRANSMEM 298  
 FT TRANSMEM 319  
 FT TRANSMEM 337  
 FT TRANSMEM 338  
 FT TRANSMEM 358  
 FT TRANSMEM 377  
 FT DISULFID 108  
 FT LIPID 377  
 FT SEQUENCE 416 AA; 46265 MW; 127FC5F5CB6FE208 CRC64;

Query Match 18.6%; Score 401; DB 1; Length 416;  
 Best Local Similarity 29.4%; Pred. No. 4.1e-19;  
 Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;

QY 18 WPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTWMLIGRYDMRT 71  
 Db 6 WPPRPSGAGSLERLGLVDTRLMAKVFETALYSILFAGTGNALSHVVLKARTGRPG 65  
 QY 72 TTNYLIGSMAYSDLLIL-GLPFDLYR-LMRSRPWFGLCLSLYVGGCTYATLLH 128





DT 01-JUN-1993 (Rel. 26, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Somatostatin receptor type 5 (SSSR).  
 GN SSTR5.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Plutitary;  
 RX MEDLINE=93125493; PubMed=1362243;  
 RA O'Carroll A.-M., Lolait S.J., Konig M., Mahan L.C.;  
 RT "Molecular cloning and expression of a pituitary somatostatin  
 RT receptor with preferential affinity for somatostatin-28.";  
 RL Mol. Pharmacol. 42:939-946(1992).  
 RN [2]  
 RP REVISIONS TO C-TERMINUS.  
 RC TISSUE=Plutitary;  
 RX MEDLINE=94195267; PubMed=7908405;  
 RA Panetta R., Greenwood M.T., Warszynska A., Demchyshyn L.L., Day R.,  
 RA Niznik H.B., Srikant C.B., Patel Y.C.;  
 RT "Molecular cloning, functional characterization, and chromosomal  
 RT localization of a human somatostatin receptor (somatostatin receptor  
 RT type 5) with preferential affinity for somatostatin-28.";  
 RL Mol. Pharmacol. 45:417-427(1994).  
 CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS  
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: PROMINENT IN THE PITUITARY AND SMALL  
 CC INTESTINE. LOW LEVELS IN ISLETS AND SPLEEN. NOT DETECTED IN  
 CC KIDNEY, PANCREAS, CEREBELLUM, OR CORTEX.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, L04535; AAA17029.1; -;  
 DR EMBL, U01152; AAC09011.1; -;  
 DR EMBL, X74828; CAAS2825.1; -;  
 DR HSSP; P02699; 1F88.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm.1; 1.1.1.1.  
 DR PRINTS; PR00237; GPCRHOODPSN.  
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.  
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Lipoprotein; Palmitate.  
 FT DOMAIN 1 35  
 FT TRANSMEM 36 63  
 FT TRANSMEM 64 73  
 FT TRANSMEM 74 99  
 FT DOMAIN 100 111  
 FT TRANSMEM 112 133  
 FT DOMAIN 134 155  
 FT TRANSMEM 156 176  
 FT TRANSMEM 177 196  
 FT TRANSMEM 197 221  
 FT DOMAIN 222 247  
 FT TRANSMEM 248 273  
 FT DOMAIN 274 283  
 FT TRANSMEM 284 308  
 FT TRANSMEM 309 363  
 FT TRANSMEM 364 418  
 FT CARBOHYD 13 13  
 FT CARBOHYD 23 23  
 FT CARBOHYD 186 186  
 FT DISULFID 110 185  
 BY SIMILARITY.

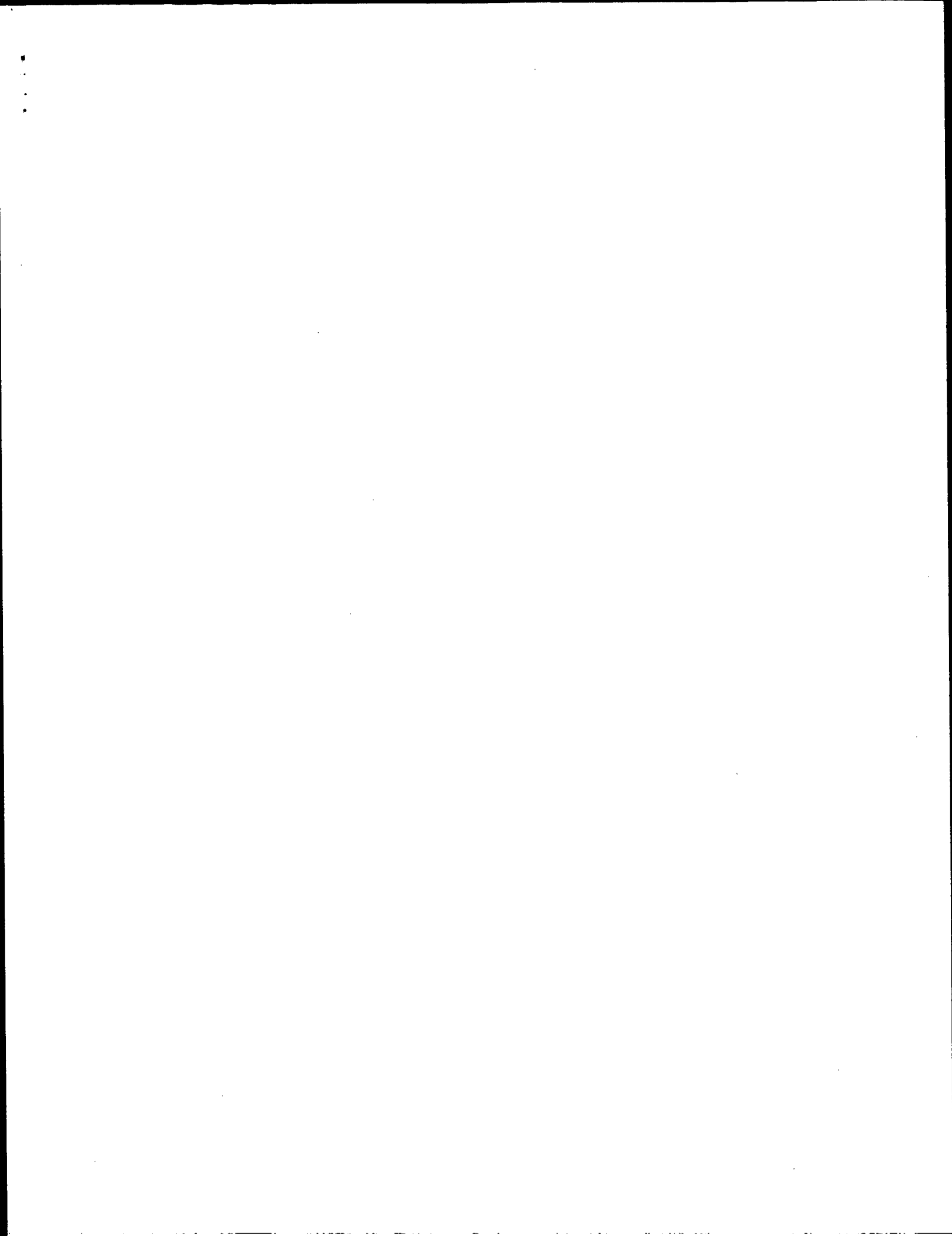
FT LIPID 320 320 PALMITATE (POTENTIAL).  
 SQ SEQUENCE 363 AA; 39971 MW; 4BD4512960613BA4 CRC64;  
 Query Match 16.6%; Score 358; DB 1; Length 363;  
 Best Local Similarity 27.3%; Pred. No. 2.1e-16;  
 Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;  
 QY 5 WNGSDGEGAEPEPPMPLPCDEKRCSPFLGA---LYPVAVCCLFVVGSGVNTVM 61  
 DB 12 WNASASSGNHN--WSLVG-----SASPWGARVLVPLVLC--TVGLSGNTLVY 60  
 QY 62 LIGYRDMRTTNLYIGSMASVDDLILGLPDIYRLRSRPWVGPLCRSLVYEGC 121  
 DB 61 VVLRHAKKKTNYIINLAVADVLFMLGFLATQNAVSGVMPGSLCRVMTLDGIN 120  
 QY 122 TYALLMTALSVRYALICPLPARVLVTRRRALIAVMVALVLSAGPLFLVGEQ 181  
 DB 121 QFTSIFCLMWASVDYRLAVVPLRSARWRBRVAKMASAAVVSFLMSLPVLVFDVQ 180  
 QY 182 DPGISVPLNGTARIASPLASSPPLMSRAPPPSPSGETMAALFSECRSPPAQ 241  
 DB 181 G-----WGTONLS-----W-----PEPVG 194  
 QY 242 LGALRWMLWTTAVFLPPLCLSLYGLI-----GRELWSSRRPLPGPAAGRGGR 294  
 DB 195 LMGAAFTYTSVIGFPGPLVVICLYLLIVKKAAGRWGSSR-----RRRSE 244  
 QY 295 QTRKVLVVLAFICLPFPVHGIIYIN--TEDSRMYFSGYFNIVLQLYLSINPI 353  
 DB 245 KVRBMVVVVVLFVGCWLPFFIVNVLAFLLPPEPTSGAGLYFVVVLS--YANSCANPL 302  
 QY 354 LVNLISKRYRAAFKLLARKSRPGRFHSRDTAGEVAGDTGG 396  
 DB 303 LVGFLSDNFRQSPKRVLCAR---RGYGMEDADAIFRRPDMSG 341  
 RESULT 14  
 ID SSR3 HUMAN STANDARD; PRT; 418 AA.  
 AC P32745;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Somatostatin receptor type 3 (SSSR) (SSR-28).  
 GN SSTR3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93149123; PubMed=1337145;  
 RA Yamada Y., Reisine T., Iaw S.F., Iliara Y., Kubota A., Kagimoto S.,  
 RA Saito M., Saito Y., Bell G.I., Saito S.;  
 RT "Somatostatin receptors, an expanding gene family: cloning and  
 RT functional characterization of human SSTR3, a protein coupled to  
 RT adenylyl cyclase.";  
 RT Mol. Endocrinol. 6:2136-2142(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93338970; PubMed=8097479;  
 RA Corness J.D., Demchyshyn L.L., Seeman P., van Tol H.H.M.,  
 RA Srikant C.B., Kent G., Patel Y.C., Niznik H.B.;  
 RT "A human somatostatin receptor (SSSR3), located on chromosome 22,  
 RT displays preferential affinity for somatostatin-14 like peptides.";  
 RL FEBS Lett. 321:279-284(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
 RA Clamp M., Smith L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Bagley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.F., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,

RA Buttrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.U., Cox A.V., Davis U., Dawson E.,  
 RA Dhami P.D., Dockree C., Dodsworth S.U., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fay J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
 RA McLaren I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,  
 RA McElay J., McLaren S., Murray A.A., Milne S.A., Mortimore B.J.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Stinson J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shitani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree U., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.T.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,  
 RA Zhang W., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,  
 RA Schet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,  
 RA Korf I., Bodel J.A., Hillier L., Mardis E., Waterston R., Wilson R.,  
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,  
 RA Mederwald H.E., Johnson A., Wong A.C.C., Morrow B.E., Edlmann L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., O'Byrne K.P.,  
 RA Seroussi E., Franzon I., Tapia I., Bruder C.E., Poyden K.P.,  
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tiliakos Y., Wright H.;  
 RA "The DNA sequence of human chromosome 22.",  
 RA Nature 402:489-495(1999).  
 CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATINS-14 AND -28. THIS RECEPTOR IS  
 CC COUPLED VIA PERTUSSIS TOXIN SENSITIVE G PROTEINS TO INHIBITION OF  
 CC ADENYLYL CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: BRAIN, PITUITARY AND PANCREAS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk).  
 CC -----  
 CC EMBL: M96738; AAA60592.1; -  
 CC EMBL: Z82188; CAB45263.1; -  
 CC PIR: S32501; S32501.  
 CC PIR: A46226; A46226.  
 CC HSSP: P34996; 1DDD.  
 CC GeneW: HGNC:11332; SSTR3.  
 CC MIM: 182453; -  
 CC InterPro: IPR000276; GPCR\_Rhodopsn.  
 CC Pfam: PF00001; 7tm\_1.1.  
 CC PRINTS: PR00237; GPCR\_Rhodopsn.  
 CC PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.  
 CC PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
 CC Multi-gene family; Polymorphism.  
 FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 44 69 1 (POTENTIAL).  
 FT DOMAIN 70 79 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 80 101 2 (POTENTIAL).  
 FT DOMAIN 102 116 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 117 138 3 (POTENTIAL).  
 FT DOMAIN 139 161 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 162 181 4 (POTENTIAL).  
 FT DOMAIN 182 205 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 206 231 5 (POTENTIAL).  
 FT DOMAIN 232 257 6 (POTENTIAL).  
 FT TRANSMEM 258 279 6 (POTENTIAL).  
 FT DOMAIN 280 293 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 294 316 7 (POTENTIAL).  
 FT DOMAIN 317 418 7 (POTENTIAL).  
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 116 191 BY SIMILARITY.  
 FT DOMAIN 346 360 S -> T (IN DBSNP:229568).  
 FT VARIANT 411 411 /FTID=VAR\_011853.  
 SQ SEQUENCE 418 AA; 45847 MW; 1227095FE01190C4 CRC64;  
 Query Match 16.5%; Score 356.5; DB 1; Length 418;  
 Best Local Similarity 26.9%; Pred. No. 2.9e-16;  
 Matches 100; Conservative 63; Mismatches 150; Indels 59; Gaps 9;  
 QY 11 PEGAREPWPALPPCDERCSPPFL-----GALVPTAVCLLFYGVGVVTVMLIGR 65  
 15 PENA-SSAMPDADLVGNVAGSPGALVSGVLPLVIVVC--VGLIGNSLIVIVLR 71  
 QY 66 YRDRTTNLYIGMAVSDLLILGLPDLVRLMRSRPVFEGLLCRLSLVYGEQTVAT 125  
 72 HTASPSVNVVITLNLADLFMLGLPF-LAQNALSWPFSIMCRLTMAVDGINQFTS 130  
 QY 126 LHMATLSVERLYACRPLRAVLTTRRRVRLVLMVALLSAGPFLFYGVEDDPEI 185  
 131 IFCVLWSDVRLVAVHPTBSARMTAPARVSAVVAASVAVVLPVVVFGSV----- 184  
 QY 186 SVFGLNTATIASPLASSPPVWLSRAPPEPSGPETAEEAALFSRECRSPAGLAL 245  
 185 -----PRGMSTCHMO-----WEPALAAWA 204  
 QY 246 RYMLVTTAVFPLPLCLSLIYGLIGRLMSRRPLRGPAASRGRRHQRKRLVLLVYL 305  
 205 GFITITALLFGFPLPLVCLCLLLLVKRSAGRWAPSCQRRRSRRVTRMVAVVA 264  
 QY 306 AFITGLDPFHVGRITTYINTE-DSRMVYSQYNIYVALQLFYLSASINPLTYNLSKKYRA 364  
 265 LFLVCMPEFYVIVNVVCPLEBPAPFGLYPLVVALP--VANSANPLLYGLSYLRFK 321  
 QY 365 AAFKLLARKSR 376  
 Db 322 QGFRVLRPSR 333  
 RESULT 15  
 ID GP39\_HUMAN STANDARD; PRT; 453 AA.  
 AC 043194;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative G protein-coupled receptor GPR39.  
 GN GPR39.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98110578; PubMed=9441746;  
 RA McKee K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,  
 RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;  
 RT "Cloning and characterization of two human G protein-coupled receptor  
 genes (GPR38 and GPR39) related to the growth hormone secretagogue  
 and neurotensin receptors."

RL Genomics 46:426-434(1997). Integral membrane protein.  
CC -1- SUBCELLULAR LOCATION: Family 1 of G-PROTEIN COUPLED RECEPTORS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF034633; AAC26082.1; -  
DR GenBank; HGNC:4496; GPR39.  
DR MIM; 602886; -  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
DR PROSITE; PS00262; G\_PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
FT DOMAIN 1 25  
FT TRANSMEM 26 46  
FT DOMAIN 47 71  
FT TRANSMEM 72 92  
FT DOMAIN 93 109  
FT TRANSMEM 110 131  
FT DOMAIN 132 154  
FT TRANSMEM 155 175  
FT DOMAIN 176 223  
FT TRANSMEM 224 243  
FT DOMAIN 244 280  
FT TRANSMEM 281 301  
FT DOMAIN 302 321  
FT TRANSMEM 322 347  
FT DOMAIN 348 428  
FT DISULFID 108 210  
FT CARBOHYD 192 192  
FT CARBOHYD 206 206  
FT CARBOHYD 212 212  
SQ SEQUENCE 453 AA; 51328 MW; 8E3A233420D9021E CRC64;  
  
Query Match 16.2%; Score 349.5; DB 1; Length 453;  
Best Local Similarity 28.5%; Pred. No. 8.8e-16; Indels 45; Gaps 13;  
Matches 111; Conservative 64; Mismatches 170;  
  
QY 20 PALPDCRRCS-----PPLGALVPT--AVCLCLFVVGSGNVTV---MLIGR 65  
DB 4 PSLPQSD---CSQIDHSHVPEFEVATWIKILIVYLIIFFWGLGN SATIRVQVLOK 60  
QY 66 YRDKRTTNLYIGSMVAVSDLI--LGLPFDLYR--LMRSRPWFGLLCRLSLYVGECTY 123  
DB 61 KGYLOKEVTDHNVSLACSDILVFLIGMPEFYSIINNPULTSSYTLSCKLHFLFEACSY 120  
QY 124 ATLLHMTALSEVRYLAI CPTLRARVLT RRRVALI AVLMAVALSAGPFLFVGEODP 183  
DB 121 ATLLHVLTLSEFRYTAICHFRYKAVSGCQVXLIGFWVTSAVALPLFLFAMGTEY-P 179  
QY 184 GISVVPGLNGTARLASSPLASSPPLMLSRAPPSPSPGETAALFSGRCPSPAQ-- 241  
DB 180 LVN--VPSHRGLTCNRSSTRHNEQ-----PETSMSICTNLSSRWTVFQSS 223  
QY 242 -LGALRWLWLTATAFFLPFLCLSLIYGLIGRELMSRRP--LRGPAASGRERGHROTQR 298  
DB 224 IFGAFVYLVLLVLSAFVCMNMVQVLMKSQKSLAGGTRPQLRKSESESESRARROTI 283  
QY 299 VLVVVAFLIICWLPFHVGRLLIY--NTEDSRMYFSQYFNIVALQ--LFYLSASINPIL 354  
DB 284 FLRLIVVTLAVCMWPNQIRITMAAKPKHDWTRSYFRAYMILLPSETFFYLSVINPIL 343  
QY 355 YNLISKYTRAAAFKLLLRKSRPGRGFRSR 384  
DB 344 YTVSSQOFRFVVOVLCCLSLQSHANHEKR 373

Search completed: March 16, 2003, 15:04:19  
Job time : 15 secs



GenCore version 5.1.4\_p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:10 ; Search time 18 Seconds  
(without alignments)  
2200.410 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156

Sequence: 1 MGSPWNGSDGEGAREPPMP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73.\*  
2: PIR1.\*  
3: PIR2.\*  
4: PIR3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	23.1	424	2 JH0164	neurotensin recept
2	491	22.8	418	2 S29506	neurotensin recept
3	434	20.1	418	2 A88013	protein K10B4.4
4	401	18.6	415	2 S68832	neurotensin recept
5	378	17.5	378	2 T15816	hypothetical prote
6	358	16.6	363	2 I57940	somatostatin recep
7	356.5	16.5	418	2 A46226	somatostatin recep
8	343	15.9	363	2 I57955	somatostatin recep
9	343	15.9	364	2 JN0763	somatostatin recep
10	342.5	15.9	380	2 A55259	kappa opioid recep
11	338	15.7	352	2 JE0296	thyrotropin releas
12	338	15.7	380	2 S36143	kappa opioid recep
13	334	15.5	380	2 JC2338	kappa opioid recep
14	334	15.5	388	2 JN0605	somatostatin recep
15	333.5	15.5	367	2 I49022	kappa opioid recep
16	333.5	15.5	367	2 JC2421	opioid receptor ho
17	333.5	15.5	367	2 I56520	G protein-coupled
18	333.5	15.5	428	2 S05058	probable G protein
19	330.5	15.3	428	2 A44021	somatostatin recep
20	329.5	15.3	370	2 S43087	orphan opioid rece
21	328	15.2	380	2 A48227	kappa opioid recep
22	327	15.2	380	2 JC2434	kappa opioid recep
23	326.5	15.1	384	2 A47249	brain-specific som
24	325.5	15.1	372	2 I38532	delta opioid recep
25	321	14.9	384	2 UC4629	somatostatin recep
26	320.5	14.9	372	2 S34592	delta opioid recep
27	320	14.8	398	2 JN0708	thyrotropin-releas
28	320	14.8	519	2 S17783	tachykinin recepto
29	319	14.8	372	2 B48227	delta opioid recep

30	318.5	14.8	387	2 JC5949	galanin receptor 2
31	318	14.7	380	2 I38435	angiotensin recept
32	317	14.7	398	1 J01059	neurokinin 2 recep
33	317	14.7	411	2 I56444	thyrotropin-relea
34	317	14.7	412	2 S23436	thyrotropin-relea
35	316	14.7	393	2 A39251	thyrotropin-relea
36	314	14.6	391	2 C41795	somatostatin recep
37	311	14.4	373	2 JE0087	delta opioid recep
38	311	14.4	384	1 S00516	neurokinin 2 recep
39	310	14.4	391	2 A39297	somatostatin recep
40	308.5	14.3	369	2 D41795	somatostatin recep
41	308.5	14.3	392	2 S65693	opioid receptor mu
42	308.5	14.3	400	2 I56553	opioid receptor mu
43	308	14.3	391	2 A41795	somatostatin recep
44	308	14.3	398	2 I56517	mu opioid receptor
45	307.5	14.3	369	2 A45291	somatostatin recep

#### ALIGNMENTS

##### RESULT 1

JH0164

neurotensin receptor - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 17-Mar-2000

C/Accession: JH0164

R/Tanaka, K.; Masu, M.; Nakanishi, S.

Neuron 4, 847-854, 1990

A/Title: Structure and functional expression of the cloned rat neurotensin receptor.

A/Reference number: JH0164; PMID:90297956; PMID:1694443

A/Accession: JH0164

A/Molecule type: mRNA

A/Residues: 1-424 <TRAN>

C/Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. Th

eter (neurotensin receptor) in the brain and as a hormone) cellular mediator in peripheral tis

C/Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:97-121/Domain: transmembrane #status predicted <TM1>

F:144-165/Domain: transmembrane #status predicted <TM2>

F:189-210/Domain: transmembrane #status predicted <TM3>

F:236-260/Domain: transmembrane #status predicted <TM4>

F:309-330/Domain: transmembrane #status predicted <TM5>

F:348-372/Domain: transmembrane #status predicted <TM6>

F:38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 23.1%; Score 498; DB 2; Length 424;

Best Local Similarity 32.9%; Pred. No. 1.1e-32;

Matches 125; Conservative 58; Mismatches 111; Indels 86; Gaps 11;

QY	39	VPTAVCLCFVVGSGNVTVWLIGR---YRDMRTTNLYIGSMAYSDLLI-LLGLPFD 94	
DB	65	VLTAVLYLAFVVGTVGNSVTAFTLARKKSLQSQSVHVLGSLASDLLILLAMFEVE 124	
QY	95	LXR-LMRSPWVGPLLCRLSLVVGEGCTATLHMTALSVERYLATCRPLARAVLVR 153	
DB	125	LVNFIWVHPMAFGDAGCGRYVFLRDACVATATLVNVSIVERTLALCHPKATLMSRS 184	
QY	154	RVALAVLMAVALLSAGPFLVGVQDPGISVVPGLNGTARTIASSPLASSPPLMUSRA 213	
DB	185	RTKFKISAIWLSALALAIPLFTMGQONRSGDTHPG-----GLVCTPIYDT----- 231	
QY	214	PPSPSPSGPETBAALAFSRECRPSPAQGLALRVMLVMTTAAVFL-FLCLSLTYGLIGR 272	
DB	232	-----ATVKAVIQVNTFMSFLPMLVTSIINTVIAN 262	
QY	273	ELWSSRRPLRGPAASGR-----ERGRQTR-----VLLVVLAFII 309	
DB	263	KLTVM---VHQAAEQGRVCTVGTGHNGLSHSTFMWTEIPGRVQALRHGVLLRANVIAFVV 319	
QY	310	CHLPFVIGRLIYINTDSR---MMTFQYFNIVAIQFLSLASINPILNLISKYRRAA 365	

```

Db      320 CMLPYNVRRLMFCYISDEQWTTFLFDFPHFYFMLTNALFYVSSAINPIILYNLVSANFRGV 379
Qy      366 AFKLIA-----RKSRP 377
           |         |||
Db      380 FLSTLAACLCPGWRHRRKKRP 399

RESULT 2
S29506
neurotensin receptor - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S29506
R:Vita, N.; Laurent, P.; Lefort, S.; Chalton, P.; Dumont, X.; Kagnad, M.; Gully, D.; Le F
RBBS Letc. 317, 139-142, 1993
A>Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
A:Reference number: S29506; MUID:93154505; PMID:8381365
A:Accession: S29506
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <VT>
A:Cross-references: EMBL:X70070, NID:g35020, PIDN:CAA49675.1, PID:g35021
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match          22.8%; Score 491; DB 2; Length 418;
Best Local Similarity 32.8%; Pred. No. 4e-32;
Matches 125; Conservative 62; Mismatches 118; Indels 76; Gaps 13;

Qy      39 VPTAVCLCFVVGVSGNVTVTMLIGR---YEDMTTTLNYGSMAVSDLL-ILLGLPRD 94
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      64 VLVAIVYLAFVVGTVGNITVFATFLARKKSLDSLOSTHYHGSLASDLTILLAMPVE 123
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      95 LYR-LMSRPWFPGPLLCRLSLVYGEGCTYTALTMTALSVERVIATCRPLRAVLVTR 153
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      124 LYNFIWHHPAFDGAGRGYYFLDACTYATNLVASLSVERIYLAICHPFKAKTLMRS 183
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      154 RVRALILVALNALISAGPFLFLVGEVDPGISVPGLNGTRIASSPLASPWLMSRA 213
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      184 RTKKEIFSAIMWASALLTVPMFLTWG-EQN-----                211
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      214 PPSPFGPEFAEAAALFSRECRPSPAOLGARVMMLWT-AFFLPFLCLSYGLIGR 272
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      212 -----KSADQHAGLVCTPT-IHTATVKVIVQNTFMSPFIRPMVISLVNIIAN 261
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      273 ELWSSRRPL-RGPAS-----GRGRHQTKRYLLVVLAFITCMLPEHV 316
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      262 KLTFVWROARAQGYCVCTTGEGESTFSMAIEPGRVQLRHGVRLAVLIAVFWCWLPYHV 321
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      317 GRITYINTESR--MMY-FSQYENIVALQLFYLSINPIILYNLSKYR-----AA 365
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      322 RLIMECYISDQMPPFLVDFHYHFYMTNALFYVSSIIPILYNLVSANFRHIPLATLAC 381
           |||||   :   :   :   :   :   :   :   :   :   :   :   :   :   :

Qy      366 AFKLLARKSRPGPHRSRD 386
           |       |||    |||
Db      382 LCPVWRRRRRKR-APSKKADS 401
           |       |||    |||

RESULT 3
A88013
protein KI0B4_4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88013
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9651916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88013
A>Status: preliminary
A:Molecule type: DNA
```

A;Residues: 1-118 <STD>  
A;Cross-references: GSP:chr II; PIDN:AB71009.1; PID:g2429475; GSPDB:GN00020; CESP:K10B4  
A>Note: similar to family I of G-protein coupled receptors  
C;Genetics:  
A;Gene: K10B4.4  
A;Map position: 2

```

Query Match          20.1%; Score 434; DB 2; Length 416;
Best Local Similarity 30.1%; Pred. No. 1.6e-27;
Matches 109; Conservative 77; Mismatches 138; Indels 38; Gaps 10;
```

Qy	29	RCSPPFLGALVPTAVCLCFVGVSGNVVTYMLICRYRDMRTTTLXGLSMAVSDDL-I	87
Dd	24	RCQ--SAGIVIPVLIIYGTIFLLGLFGNICITCVIANANKSMNPVTYYFSLAVSDIIL	81
Qy	88	LIGLPEPDLYR-LMRSPWVEGFLLCRLSLVYGEGCYATLLHNTALSVERYLACIRPLRA	146
Dd	82	IIGLPDEFQSIDYSPIYKFSEGCARAFLEIFTSYASIMITCCSFEEWLAIQHPLRS	141
Qy	147	RVLVTRRRRALIAVAAMVALISAGPFLPLVGVEDP--GISVPGLNGTARIASSPLA	203
Dd	142	KIFSTLMRANVILLIAMTISFWCALPIAFVIQINKLPLEDAKYQEPWTNKVSFFAVGVLN	201
Qy	204	SSEPPLWSRAPPSGPSPTAEAAALFSREC-----RSPADLGALRWLMTTAFFEL	258
Dd	202	-----NRIFPVSS-----TDGI FVLHTEFCMKNOSRPDQQK---IIIFAFTVEFI	244
Qy	259	PFLCLSLVLGLIGRELMSSRRPLRG--AASGRERGHQRKRVLVVLAIFIQLPEFHV	316
Dd	245	PAIAIYMAHAIAVQESSEIDLDGDGMVKKRHKNSRIYLLKMLSVITFFICMLPEHI	304
Qy	317	GRIIVINTEDSRMVMYFSQYFNIVALQLFYLS-----ASINPILYNLIKRYRAAAFKL	369
Dd	305	QRLSYTTTWSETTTISPPOFLSMIVFYISGHCYYSNAANPILNYILSQKXSAPFCRT	364
Qy	370	LT 371	
Dd	365	IL 366	

RESULT 4  
S68822  
neurotensin receptor 2, leucobastine-sensitive - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Dec-1997 #sequence\_reviseun\_12-Dec-1997 #ext\_change 20-Jun-2000  
C;Accession: S68822  
R;Chalon, P.; Viltz, N.; Kaghad, M.; Guillemot, M.; Bonnin, J.; Delpech, B.; Le Fur, G.  
FEBS Lett. 366, 91-94, 1995  
A;Title: Molecular cloning of a leucobastine-sensitive neurotensin binding site.  
A;Reference number: S68822; MUID:96228041; PMID:8647296  
A;Accession: S68822  
A;Molecule type: mRNA  
A;Residues: 1-416 <CHA>  
A;Cross-references: GB:X97121; NID:g1483579; PIDN:CMA65787.1; PID:g1483580  
A;Experimental source: hypothalamus  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: G protein-coupled receptor; transmembrane protein  
F;33-58/Domain: transmembrane #status predicted <TM1>  
F;70-91/Domain: transmembrane #status predicted <TM2>  
F;110-131/Domain: transmembrane #status predicted <TM3>  
F;155-175/Domain: transmembrane #status predicted <TM4>  
F;204-230/Domain: transmembrane #status predicted <TM5>  
F;296-315/Domain: transmembrane #status predicted <TM6>  
F;335-361/Domain: transmembrane #status predicted <TM7>

```

Query Match          18.6%; Score 401; DB 2; Length 416;
Best Local Similarity 29.4%; Pred. No. 7.5e-25;
Matches 122; Conservative 60; Mismatches 145; Indels 80; Gaps 13;
```

Qy	18	PWPALP-----PCDERCSRPFGLGALVPYTVAVCLCFVGVSGNVVTYMLICRYRDMDRT	71
Dd	6	PWPRPSPSAGLSLEARLGVDTRLMARVLFYALVSLIFAFGTAGNVAHVHLKARAGR	65

QY 72 -TTNLYGSMAYSDLLILL-GLPFDLYR-LMRSRPWFGPLCRSLVYEGCTYATLH 128  
 Db 66 GLRHYVISTLASALLLLVSMPELMELNFWMSHYPMVFGDLGRGYFRELCAVATVLS 125  
 QY 129 MTALSVERIALICRPLRARVLVTRRRVRLIALVMAVALLSAGPFLVGV----- 179  
 Db 126 VASLSAEKCIACQQLRARRLITPRTRRLSLVWVASIGLALPMAVINQKHEVSADG 185  
 QY 180 EODPGISVVPGLNGTARIA-----SSPLASSPPLMSRAPPPSPGPETAEEA 228  
 Db 186 EPEPASRVCTVLSRATLQVFIQVNVLVSEFALPLATFL-----NITVNHLM 234  
 QY 229 ALFSRECPSPQALARMMLWTTATYFPLPFLCSTLYGLGRELMSSRRPLRGPAASG 288  
 Db 235 ALYS-QVPSASAQVSSISPRLELSE-----EGLLGFITWRKTLISGVQASIV 281  
 QY 289 REGRHGRK-----RVLLVVVLAFLICWLPPHVGRIITYNED-----SRMWYEQYENI 338  
 Db 282 RHKQASQIRSLQHSQVLRATVAVVTCMLPYHARLMICYIPDDGWTMLDYFHYFYV 341  
 QY 339 VALQIFYLASINPILYMLISKRYAAAFKLLARKSRPRGPHRSRDTAGEVAGD 393  
 Db 342 VTNTLFYSSAVTPILYNAVSSFR-----KTLFL-----ESLGSLSGCE 379

## RESULT 5

T15816  
 hypothetical protein C48C5.1 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
 C/Accession: T15816  
 R/Availlo, A.  
 submitted to the EMBL Data Library, November 1995  
 A/Description: The sequence of C. elegans cosmid C48C5.  
 A/Reference number: Z18410  
 A/Accession: T15816  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-378 <FAV>  
 A/Cross-references: EMBL:U39994; NID:q1055102; PID:q1055105; PIDN:AA837017.1; GSPDB:GN00  
 A/Experimental source: stratin Bristol N2; clone C48C5  
 C/Genetics:  
 A/Genes: C48C5.1  
 A/Map position: X  
 A/Intons: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1  
 C/Superfamily: adenosine receptor A1

Query Match 17.5%; Score 378; DB 2; Length 378;  
 Best Local Similarity 26.7%; Pred. No. 4.9e-23;  
 Matches 97; Conservative 70; Mismatches 135; Indels 36; Gaps 9;

QY 38 LVPYAVCLCFVVGSGNVVTMLIGRYDMRTTNLYGSMAYSDLLIL-LGLPFDLY 96  
 Db 51 LYKLTALYIFLIVGIGNTTCLVMKHPMKTHASMTLMLAVSDVTLVGLPFEVM 110  
 QY 97 RLWRSRPWFGPLCRSLVYEGCTYATLHMTALSVERYLAICRPL-RARVLVTRRV 155  
 Db 111 MNMQYPMFPEDYICNLKALIAETSSVSILITLIFAIERVAVACHPLMKVQPFKRI 170  
 QY 156 RALIAVMAVALLSAGPFLVGVQDPGISVVPGLNGARARASSPPLMSRAP 215  
 Db 171 GTIIGFTWISILCAMP--FAIHRADYIMKSWPGTDRIPIKSSKMCN----- 217  
 QY 216 PSPSPGETAAMALFSRECPSPQALGLRVL--WTTAVFPLPFLCSTLYGLGREL 274  
 Db 218 -----IAYMF-----EPKLASTFKILFHFSALAFALPLFTVILYAIACV 260  
 QY 275 WSRRLPLRGPAASRGHGRQTKVLLVLAFTICMLPHEVGRITVINTEDSRMA-YFS 333  
 Db 261 -SSRRITQPGELDITEELQKRMIMILCAIVSAFFICYLPFOQLRLLFFYDNEVILLTWN 319  
 QY 334 QYENIVALQIFYLASINPILYMLISKRYAAAFKLL 371

Db 320 QYMFISGLFYLATIIINPIAVMLASSRFR-RAFKDIL 356

## RESULT 6

I57940

somatostatin receptor 5 - rat  
 N/Alternate names: somatostatin release-inhibiting factor subtype 28 receptor  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 24-Nov-1999

C/Accession: I57940; I57949; S39244  
 R/O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.  
 Mol. Pharmacol. 42, 939-946, 1992

A/Title: Molecular cloning and expression of a pituitary somatostatin receptor with pre  
 A/Reference number: I57940; MUID:93125499; PMID:1362243  
 A/Accession: I57940

A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-363 <OCA1>

A/Cross-references: GB:I04535; NID:q409238; PIDN:AA17029.1; PID:q409239  
 R/O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.  
 Mol. Pharmacol. 44, 1278, 1993

A/Title: Molecular cloning and expression of a pituitary somatostatin receptor with pre  
 A/Reference number: I57949; MUID:94088493; PMID:8264565  
 A/Accession: I57949

A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 341-363 <OCA2>

A/Cross-references: GB:I67370; NID:q455947; PIDN:AA82371.1; PID:q455948  
 R/Penetta, R.; Greenwood, M.; Patel, Y.C.  
 submitted to the EMBL Data Library, August 1993

A/Description: Correction of the nucleotide and amino acid sequence of the rat somatost  
 A/Reference number: S39244  
 A/Accession: S39244

A/Molecule type: mRNA  
 A/Residues: 309-363 <PEN>  
 A/Cross-references: EMBL:X74828; NID:q433911; PIDN:CAA52825.1; PID:q433912  
 C/Genetics:  
 A/Genes: SSTRS  
 C/Superfamily: vertebrate rhodopsin

Query Match 16.6%; Score 358; DB 2; Length 363;  
 Best Local Similarity 27.3%; Pred. No. 1.9e-21;  
 Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

QY 5 WNSDDEPGEARPEPPWALPPCDERRCSPFLGA---LVPTAVCLCFVVGSGNVVTM 61  
 Db 12 WNSAASGNNH--WSLVG-----SASPWGAPAVLVPLVYLVC--TVGLSGNTLVIV 60  
 QY 62 LIGRYDMRTTNLYGSMAYSDLLILLGLPFDLYRLMRSRPWFGPLCRSLVYEGGC 121  
 Db 61 VVLRHAKMTVNVVILNLAADVLEMLGIPLATQAAVVSYPFGSFLCRVWTLGGIN 120  
 QY 122 TYATLHMTALSVERYLAICRPLRARVLVTRRRVRLIALVMAVALLSAGPFLVGEQ 181  
 Db 121 QFTSIFCLMWSYDRYLAVHPRLRSARWRPRVAKMAAIVWVFSILMSIPLVFAVQ 180  
 QY 182 DEGISVVPGLNGARARASSPPLMSRAPPPSPSPETAAMALSRRECRPBAQ 241  
 Db 181 G-----WGTCNLS-----W-----PPVVG 194  
 QY 242 LGALRVMLWTTATYFPLPFLCSTLYGLI-----GRELMSSRRPLRGPAASRGHGR 294  
 Db 195 LMGAAITTSVGLGFGPPLVITCLVLLIVKVKRAAGMRGSSR-----RRSEP 244  
 QY 295 QTKRVLLVVVLAFTICMLPHEVGRITVINTEDSRMAYEQYENIVALQIFYLASINPI 353  
 Db 245 KYTRVVVVVLAFFVVGCMPEFFIVNINLAFTLPEEPTSAQLYFFVVLS--YANSCANPL 302  
 QY 354 LVNLYSKRYRAAFKLLARKSRPRGPHRSRDTAGEVAGDTGG 396  
 Db 303 LVGFLSDNFRQSPFKVLCR---RGYGMEDADALIEPRDKSG 341





C:Comment: This protein is a member of somatostatin receptor family.  
 C:Genetics:  
 A:Gene: GDB:SSTR5  
 A:Cross-references: GDB:138452; OMIM:182455  
 A:Map position: 16p13.3-16p13.3  
 A:Introns: #status absent  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; phosphoprotein; cholo  
 F/40-66/Domain: transmembrane #status predicted <TM1>  
 F/70-102/Domain: transmembrane #status predicted <TM2>  
 F/114-135/Domain: transmembrane #status predicted <TM3>  
 F/155-177/Domain: transmembrane #status predicted <TM4>  
 F/196-228/Domain: transmembrane #status predicted <TM5>  
 F/246-273/Domain: transmembrane #status predicted <TM6>  
 F/280-307/Domain: transmembrane #status predicted <TM7>  
 F/312-186/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/312-186/Disulfide bonds: #status predicted  
 F/242-325/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pr  
 F/247/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predic  
 F/320/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.9%; Score 343; DB 2; Length 364;  
 Best Local Similarity 27.7%; Pred. No. 3.2e-20;  
 Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;

QY 14 AREPPPALP-----CDERR-CSPFP-LGA--LVPTAVCLCLFVVGSGVVTVM 62  
 DB 7 ASTPSWMASSPGAASGGGDNRTLVGPASAGARALVLPVLYLLVC--AAGLGNTLVIV 64  
 QY 63 IGRYDRRTTNLYGMAVSDLLILGLPFDLYRLMSRPVPRPLCRSLVYGECT 122  
 DB 65 VLRAKAKTKVNYILNLAVDLVLMGLP-LATONASFPFGPVLCRLVMTLDGVNQ 123  
 QY 123 YATLHMTALSVRYLAICRPLARVLTTRRRVRLAVLMAVALISGPFPLVGEOD 182  
 DB 124 FTSVFCULTVMSVDYIALVHPLSARMRPRVAKLASAAWLSLMSIPLVPADVC- 182  
 QY 183 PGISVVDGLNGTARIASSPLASSPPLWLSRAPPSGPGETAABAAALFSRECRSPACL 242  
 DB 183 -----GGTCNAS-----W-----DEPVGL 196  
 QY 243 GALRVMLVTTAVFPLPCLSLIYGLIGRELMSRRPLRGPASG-----RENGHROT 296  
 DB 197 WGAFTITYTAVLGFAPLVLVLCCLLVIVK-----RAAGVRVGCYRRRSERV 246  
 QY 297 KRVLLVVLVAFIICMLPFHVGRIIYI-----NTEDSRMYFSQYFNIVALQLFYLASIN 351  
 DB 247 TRWLVVVLVFPAGCMLPEFTVNYVLAVALQEPASAGLIF--FVVI--LSYANSCAN 300  
 QY 352 PLYNLISKYRRAAFKLLARK 374  
 DB 301 PVLVGFUSDNRQSFQVCLARK 323

RESULT 10  
 A55259  
 kappa opioid receptor - guinea pig  
 N:Alternate names: dynorphin receptor  
 C:Species: Cavia porcellus (guinea pig)  
 C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 24-Nov-1999  
 C:Accession: A55259  
 R:Ale, G.; Meng, F.; Mansour, A.; Thompson, R.C.; Hoversten, M.T.; Goldstein, A.; Watson  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3779-3783, 1994  
 A>Title: Primary structure and functional expression of a guinea pig kappa opioid (dynor  
 A:Reference number: A55259; MUID:94224825; PMID:8170987  
 A:Accession: A55259  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <XIE>  
 A:Cross-references: GB:U04092; NID:9476106; PIDN:AAA67171.1; PID:9476107  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: transmembrane protein

Query Match 15.9%; Score 342.5; DB 2; Length 380;  
 Best Local Similarity 25.1%; Pred. No. 3.6e-20;  
 Matches 102; Conservative 74; Mismatches 147; Indels 83; Gaps 12;

QY 2 GSPW-----NGSGPFGAREPP--WPAUPCCEBRCRSPPLGALVVTAVCLCF 49  
 DB 26 GSAMLPQWABPDGSGAGPQDEQLPEAHISPAIP-----VIITVYSVVF 70  
 QY 50 VVSGVNVVTVMLIGRRDRRTTNLYGMAVSDLLILGLPFDLYRLMSRPVPGPL 109  
 DB 71 VVGLVGNLSLVMPFIITRTKTKATNIIYPLALADALVTTMPQ-STVYLMNSPFEDV 129  
 QY 110 LCRSLSYVEGCTYATLLHMTALSVRYLAICRPLARVLTTRRRVRLAVLMAVALS 169  
 DB 130 LCKIVISIDVYNNFTSIFLTVMSSVDYIAVCHVKALDPRTPKAKIINICIVLSSSV 189  
 QY 170 AGFPLVLYVEQDPGISVVDGLNGTARIASSPLASSPPLWLSRAPPSGPGETAABAA 229  
 DB 190 GISAILLGTKREDDVITIE-----CSLQFPDDDSWWD----- 223  
 QY 230 LFSRECRPPPAQALRVMLVTTAVFPLPCLSLIYGLIGRELMSRRPLRGPASGR 289  
 DB 224 LFWKIC-----VVPAPVIVLIIIVCTYTMILRL-KSVLLSG--SREK 265  
 QY 290 ERGHRQTKRVLLVLAFLIICMLPFHVGRIIYI--NTEDSRMYFSQYFNIVALQLFYL 347  
 DB 266 DRLRLRITRLVVLVAVFIIICWTPRIHIFLVEALGSTSHSTAALSSYF--CIALGYTN 322  
 QY 348 ASINPLILNLISKYRAA-----AFKLLAKSRPFGFRSRDTA 387  
 DB 323 SSILNPILYAFLENFRCPFRDPCFPIMKMEROSTSRVNTVDDPA 368

RESULT 11  
 J0296  
 thyrotropin releasing hormone receptor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jul-2000  
 C:Accession: J0296  
 R:Itada, H.; Nakamura, T.; Itob, J.; Iwasa, H.; Kanatani, A.; Borkowski, J.; Ihara,  
 Biochem. Biophys. Res. Commun. 250, 68-71, 1999  
 A>Title: Cloning and characterization of a new subtype of thyrotropin-releasing hormone  
 A:Reference number: J0296; MUID:98407892; PMID:9735333  
 A:Accession: J0296  
 A:Molecule type: mRNA  
 A:Residues: 1-352 <ITTA>  
 A:Cross-references: DDBJ:AB015645; NID:93660553; PIDN:BA33437.1; PID:93660554  
 C:Superfamily: adenosine receptor A1  
 F/26-48/Domain: transmembrane #status predicted <TM1>  
 F/58-80/Domain: transmembrane #status predicted <TM2>  
 F/97-118/Domain: transmembrane #status predicted <TM3>  
 F/142-165/Domain: transmembrane #status predicted <TM4>  
 F/188-209/Domain: transmembrane #status predicted <TM5>  
 F/252-273/Domain: transmembrane #status predicted <TM6>  
 F/282-304/Domain: transmembrane #status predicted <TM7>

Query Match 15.7%; Score 338; DB 2; Length 352;  
 Best Local Similarity 28.8%; Pred. No. 7.7e-20;  
 Matches 108; Conservative 52; Mismatches 109; Indels 106; Gaps 13;

QY 41 VTAVCLCLFV--VGSVGNVTVMLIGRYDMRTTNLYGMAVSDLLIL--GLPFDLY 96  
 DB 23 VVSFVLVLLVCTLGIGNMVILVLTSDMRTPTCYLSLALADLVLLAAGLPVSD 82  
 QY 97 RLMRSRPWFEGPLLCRLSYVEGCTYATLL-----HMTALSVRYLAICRPLARVL 149  
 DB 83 SL--VGHMTVGRAGCL-----GLTYFOYLGINVSFSLIAFTVERXYLAICPLRAQTV 133  
 QY 150 VTRRRVRLAVLMAVALISGPFPLVGV-----EODPGISVVDGLNGTARIASSPLA 203  
 DB 134 CVVAAKRIIAIKWGTISYCLMFLVLDVLRNDNRLECGYVPPGL----- 181  
 QY 204 SSPPLWLSRAPPSGPGETAABAAALFSRECRPPPAQALRVMLVTTAVFPLPCL 263

Db 182 -----ELMSRRPLRG--PAASGR-----EKGHQTKRVLLVVLAFI 308  
 QY 264 SILYGLIGR-----ELMSRRPLRG--PAASGR-----EKGHQTKRVLLVVLAFI 308  
 Db 202 LVLYGILIGRILFOSPLSQAQWOKEROPHQSSEAPGNCRAKSSRQATMLLVVLLFA 261  
 QY 309 ICWLPFHVGRITTYNTEDESRMWFYFSQYFNIVLQLF-----YLSASINPILYNLISKYR 363  
 Db 262 VLMTPERT--LVLLNS-----FVAQPLDPVLLPFCRTCVYINSANVAVVYSLMSQKFR 313  
 QY 364 AAAPKLLARKSRPR 378  
 Db 314 AAFLKLCWCRAAGPQ 328  
 RESULT 12  
 S36143  
 kappa opioid receptor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 10-Dec-1993 #sequence\_revision 19-Oct-1995 #text\_change 20-Jun-2000  
 C:Accession: S36143; S38825; S36102; S39015; A48789  
 R:Nishi, M.; Takeshima, H.; Fukuda, K.; Kato, S.; Mori, K.  
 FEBS Lett. 330, 77-80, 1993  
 A>Title: cDNA cloning and pharmacological characterization of an opioid receptor with hi  
 A:Reference number: S36143; MUID:93380575; PMID:8396539  
 A:Accession: S36143  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <NIS>  
 A:Cross-references: GB:D16534; NID:G409390; PIDN:BA003971.1; PID:G415310  
 R:Chen, Y.; Mestek, A.; Liu, J.; Yu, L.  
 Biochem. J. 295, 625-628, 1993  
 A>Title: Molecular cloning of a rat kappa opioid receptor reveals sequence similarities  
 A:Reference number: S38825; MUID:94059008; PMID:8240267  
 A:Accession: S38825  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <CHE>  
 A:Cross-references: GB:L22001; NID:G409236; PIDN:AAA41495.1; PID:G409237  
 R:Minami, M.; Toya, T.; Katano, Y.; Nakamura, S.; Onogi, T.; Kaneko, S.; Sat  
 FEBS Lett. 329, 291-295, 1993  
 A>Title: Cloning and expression of a cDNA for the rat kappa-opioid receptor.  
 A:Reference number: S36102; MUID:93374033; PMID:8103466  
 A:Accession: S36102  
 A:Molecule type: mRNA  
 A:Residues: 1-41, 'L', 43-380 <MIN>  
 A:Cross-references: GB:D16829; NID:G404115; PIDN:BA04109.1; PID:G404116  
 R:Li, S.; Zhu, J.; Chen, C.; Chen, Y.W.; Dertiel, J.K.; Ashby, B.; Liu-Chen, L.Y.  
 Biochem. J. 295, 629-633, 1993  
 A>Title: Molecular cloning and expression of a rat kappa opioid receptor.  
 A:Reference number: S39015; MUID:94059009; PMID:8240268  
 A:Accession: S39015  
 A:Molecule type: mRNA  
 A:Residues: 1-344, 'Y', 346-380 <LIS>  
 A:Cross-references: GB:L22536; NID:G425188; PIDN:AAA41496.1; PID:G425189  
 R:Meng, F.; Xie, G.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 9954-9958, 1993  
 A>Title: Cloning and pharmacological characterization of a rat kappa opioid receptor.  
 A:Reference number: A48789; MUID:94052210; PMID:8224341  
 A:Accession: A48789  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <RES>  
 A:Cross-references: EMBL:U00442; NID:G403486; PIDN:AAA18261.1; PID:G403487  
 C:Superfamily: Vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; transmembrane protein  
 Query Match 15.7%; Score 338; DB 2; Length 380;  
 Best Local Similarity 25.5%; Pred. No. 8,4e-20;  
 Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;  
 QY 6 NSGSDPEGARPP--WPALPPCDERRCSPPFGLVPTAVCLCLFVAVGSGNVVTWML 62

Db 39 NSGVSSEDQLEPAHISPAIP-----VITAVSVFVFGVLGNSLVNFV 83  
 QY 63 IGRYDRRTTNTNLYIGMAVSDLLILGLPDLRLWRSRBPVGPULLCRSLVYEGCT 122  
 Db 84 IIRYTKTKATNTIYIFMLALADLVTTMPQ--SAVILMSWPGDVLCKIVISIDYVM 142  
 QY 123 YATLLHMTALSVERYLAIICRPLRARVLVTRRRVALIAVLAVALLSAGPFLVGVGD 182  
 Db 143 FTSIFLTLMNSVDRIYAVCHPVKALDPRTPKAKIINICIMLLASSVSAISAVLGSTYR 202  
 QY 183 PGISVVGINGTARLASPLASSPPLMLSRAPPSPGPFTAAEALFSECRSPAPQL 242  
 Db 203 EDVDVIE-----CSLQPPDEYSWD-----LFMKIC----- 229  
 QY 243 GALKRVLMTVTAAPFLPFLCLSYGLIGRELMSRRPLRGPAASGRGRHQRVLLV 302  
 Db 230 -----VPYFAVIVPVLIIVCYTLMLTLR--KSVRLSG--SREKDRNLRIKVLV 278  
 QY 303 VLAFFICWLPFHVGRITTY--NTEDESRMWFYFSQYFNIVLQLFYLSASINPILYNLISK 360  
 Db 279 VVAFFIICWLPFHIFILVEALGSTSHSTAVLSYF--CIALGYTNSLNPLVLAFLDE 335  
 QY 361 KYRAA-----AFKLLARKSRPRGFHRSRDPAGVAG--DTGG 396  
 Db 336 NFKKCFDFCFPIKWREROST---NRKNTVQDPASMRDVG 375  
 RESULT 13  
 JC2338  
 kappa opioid receptor 1 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 19-May-2000  
 C:Accession: JC2338; A55354; I57005; G01546  
 R:Manson, B.; Bare, L.; Yang, D.  
 Biochem. Biophys. Res. Commun. 202, 1431-1437, 1994  
 A>Title: Isolation of a human kappa opioid receptor cDNA from placenta.  
 A:Reference number: JC2338; MUID:94338360; PMID:8060324  
 A:Accession: JC2338  
 A:Molecule type: mRNA  
 A:Residues: 1-380 <MAN>  
 A:Cross-references: GB:U1053; NID:G532059; PIDN:AAA20985.1; PID:G532060  
 A:Experimental source: Placenta  
 R:Wang, J.B.; Johnson, P.S.; Wu, J.M.; Wang, W.F.; Uhl, G.R.  
 J. Biol. Chem. 269, 25966-25969, 1994  
 A>Title: Human kappa opiate receptor second extracellular loop elevates dynorphin's aff  
 A:Reference number: A55354; MUID:95014415; PMID:7929306  
 A:Accession: A55354  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 136-279 <MAN>  
 A:Cross-references: GB:L3130; NID:G598184; PIDN:AAA63646.1; PID:G598185  
 R:Zhu, J.; Chen, C.; Xue, J.  
 Life Sci. 56, 201-207, 1995  
 A>Title: Cloning of a human kappa opioid receptor from the brain.  
 A:Reference number: I57005  
 A:Accession: I57005  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1, 'D', 3-380 <ZHU>  
 A:Cross-references: GB:L37362; NID:G722617; PIDN:AAA63906.1; PID:G722618  
 R:Grandy, D.K.  
 submitted to the EMBL Data Library, November 1994  
 A:Reference number: G07718  
 A:Accession: G01546  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 132-203 <GRA>  
 A:Cross-references: EMBL:U16860; NID:G595932; PIDN:AAA56758.1; PID:G595933  
 C:Comment: This receptor preferentially binds dynorphins.  
 C:Genetics:  
 A:Gene: GDB:OPR1; KOR  
 A:Cross-references: GDB:132651; OMIM:165196

A:Map position: 8q11.2-8q11.2  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; receptor; transmembrane protein  
 F:60-85/Domain: transmembrane #status predicted <TM1>  
 F:95-114/Domain: transmembrane #status predicted <TM2>  
 F:133-154/Domain: transmembrane #status predicted <TM3>  
 F:177-199/Domain: transmembrane #status predicted <TM4>  
 F:228-251/Domain: transmembrane #status predicted <TM5>  
 F:275-296/Domain: transmembrane #status predicted <TM6>  
 F:311-333/Domain: transmembrane #status predicted <TM7>

Query Match 15.5%; Score 334; DB 2; Length 380;  
 Best Local Similarity 25.2%; Pred. No. 1.8e-19;  
 Matches 99; Conservative 74; Mismatches 146; Indels 74; Gaps 11;

QY 6 NGSDGPEGAREPP---WPAIPDCERRCSPFPLGALVPYAVCLCFVVGSGNVTVML 62  
 Db 39 NGASGEQAQLEPAHISPAIP-----VITAVSVFVVGVLVGNLSMFV 83

QY 63 IGRYDMRTTNLYGMAVSDLLILGLPDLYLMSRPWFGLLCLSLYGEGET 122  
 Db 84 IIRYKMTATVIYIFNLADALVTTMPFQ-STVYLMNSWPFGLVLCXIVISIDVYM 142

QY 123 YATLHMTALSEVRYLACRPLBARVLTTRRRALAVLMAVALLSAGFLFLVGEQD 182  
 Db 143 FTSITTLTMSVDRYIAVCHPVKALDFRTPLKAKITNICWILSSVGSIAVLGGTKR 202

QY 183 PGISVPLGNTARIASPELASSPPLMLSRAPPSPSGETAAMALFSRECRPSPAQL 242  
 Db 203 EDVPIE-----CSIQFPDDYSWMD-----LPMKIC----- 229

QY 243 GALKRMLVWTVYFPLPCLSLYLGLIRELMSRRPLRGPAASRGRRHGTQKVLV 302  
 Db 230 -----VFIFPAFVLPVLIIVCYLMLRL-KSVRLLSG--SRKDNRLRITRLVLV 278

QY 303 VVLAFLICMLPEHVGRIYI-NTEDSRMVFQYFNIVALOLFYSASINPLLYNLISK 360  
 Db 279 VVAIVFVCMTPHIFILVEALGSTSHSTAALSSYF---CIALGYTSSLNPLVAFLE 335

QY 361 KYRAA-----AFKLLARKSRPRGFHRSRDTA 387  
 Db 336 NFKRCFRDPCFPLKVMEROSTSRVNTVQDPA 368

RESULT 14  
 JN0605  
 somatostatin receptor 4 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text\_change 20-Jun-2000  
 C/Accession: JN0605; JN0762; A47457  
 R;Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.  
 Biochem. Biophys. Res. Commun. 193, 648-652, 1993  
 A>Title: Molecular cloning and sequencing of a human somatostatin receptor, hsstr4.  
 A:Reference number: JN0605; MUID:93290656; PMID:8515564  
 A:Accession: JN0605  
 A:Molecule type: DNA  
 A:Residues: 1-388 <XUY>  
 A/Cross-references: GB:U4856; NID:G292499; PIDN:AAA6623.1; PID:G292500  
 R;Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I  
 Biochem. Biophys. Res. Commun. 195, 844-852, 1993  
 A>Title: Cloning, functional expression and pharmacological characterization of a fourth  
 A:Reference number: JN0762; MUID:93384611; PMID:8373420  
 A:Accession: JN0762  
 A:Molecule type: DNA  
 A:Residues: 1-388 <YMA>  
 A/Cross-references: GB:U16826; NID:G693907; PIDN:BA04106.1; PID:G693908  
 R;Hohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schulte, R.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993  
 A>Title: Cloning and characterization of a fourth human somatostatin receptor.  
 A:Reference number: A47457; MUID:93248256; PMID:8483934  
 A:Accession: A47457  
 A:Molecule type: DNA  
 A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>

A/Cross-references: GB:U07833; NID:G307429; PIDN:AAA60565.1; PID:G307430  
 A>Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:P:130858)  
 C:Comment: This protein mediates the diverse actions of the tetradecapeptide somatostatin  
 C:Genetics:  
 A:Gene: GDB:SSTR4  
 A/Cross-references: GDB:202662; OMIM:182454  
 A:Map position: 20p11.2-20p11.2  
 A:Introns: #status absent  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; ph  
 F:47-73/Domain: transmembrane #status predicted <TM1>  
 F:84-109/Domain: transmembrane #status predicted <TM2>  
 F:111-142/Domain: transmembrane #status predicted <TM3>  
 F:162-184/Domain: transmembrane #status predicted <TM4>  
 F:208-238/Domain: transmembrane #status predicted <TM5>  
 F:257-284/Domain: transmembrane #status predicted <TM6>  
 F:291-314/Domain: transmembrane #status predicted <TM7>  
 F:324/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F:319-198/Disulfide bonds: #status predicted  
 F:161,253/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status  
 F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.5%; Score 334; DB 2; Length 388;  
 Best Local Similarity 26.9%; Pred. No. 1.8e-19;  
 Matches 105; Conservative 65; Mismatches 128; Indels 92; Gaps 15;

QY 1 MGSPW---NGSDGPEGAREPPWPAIPDCERRCSPFPLGALVPYAVCLCFVVGSGN 56  
 Db 15 LGTAMPAAANASSAPABAE---AVAGBDAR-----AAGNVAIQCIYALCVGLVGN 65

QY 57 VTVMLIGRYDMRTTNLYGMAVSDLLILGLPDLYLMSRPWFGLLCLSLY 116  
 Db 66 ALVIFVILRYAKMTATNIYILNLAVALDELPMLSVF-VASSAALRHMPFGSVLCRAVL 124

QY 117 VGBGCTATLHMTALSEVRYLACRPLBARVLTTRR-VRALIV-LMAVALLSAGF 173  
 Db 125 VDGLMFTSVFCLTVLSVDRVAVHPRA--ATRRPSVAKLINIGWLASILLVLP 181

QY 174 LFLVGEQDPSISVPLGNTARIASPELASSPPLMLSRAPPSPSGETAAMALPSR 233  
 Db 182 AIFADTR-----PARGOA--VACNLQWPHAW-----SAVF-- 211

QY 234 ECRPSPAQLGALKRMLVWTVYFPLPCLSLYLGLIREL-----WSSRRLRGA 285  
 Db 212 -----VVYTFPLGLFPLVLAIGLCYLLVIGKRAVALRAGWQOR----- 251

QY 286 ASGRGRHGTQKVLVAVLAFICMLPEHVGRI--IYINEDSRMVFQYFNIVALQL 343  
 Db 252 -----NSEKITRLVLMVVVFPVLCMPFYVQQLNLVVTSLDATV-----NHVSLIL 299

QY 344 FYLSASINPLLYNLISKRYRAAFKLLAR 373  
 Db 300 SYANSCANPLLYGFLSDNFRSRQRVLCR 329

RESULT 15  
 I49022  
 kappa opioid receptor 3 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text\_change 19-May-2000  
 C/Accession: I49022  
 R;Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.I.; D  
 Mol. Pharmacol. 47, 1180-1188, 1995  
 A>Title: Cloning and functional characterization through antisense mapping of a kappa 3  
 A:Reference number: I49022; MUID:95327076; PMID:7603458  
 A:Accession: I49022  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-367 <RES>  
 A/Cross-references: EMBL:U09421; NID:G551484; PIDN:AAA8133.1; PID:G551485  
 C:Superfamily: vertebrate rhodopsin

Query Match 15.5%; Score 333.5; DB 2; Length 367;

Best Local Similarity 27.1%; Pred. No. 1.9e-19;  
Matches 101; Conservative 67; Mismatches 132; Indels 73; Gaps 12;

```
QY 34 ELGALVPTAVCLCLEVVGVSQNVVTMLIGRYRDMRTTNLYLSMAVSDLLILGLPF 93
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 44 FLGLKVTIVGLYLAICIGLLGNCLVMYVILRHRTKMKKTATNIIYFNIALADTLVLLTLP 103
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 94 ---DIYRLMRSPWVPGPLLCRLSLVVGEGCTYATLLHMTLSVERYLAICRPLRARVLV 150
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 104 QGTDLILGF---WPGNALCKIVIAIDYYNMFTSTFTLTAMSDRYVAICHPIRALDVR 159
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 TFRRYRALIAVMAVALLSAGPFLVGVGEQDPGISVPGINGTARIASSPLASSPPLWL 210
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 TSSKQAQVAVVAIMWALASV-----VGVPVAIMGSAQVEDEIEC-----L 198
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 SRAPPPSPSPGPEETAALFSEKCRPSPAQLGALAVMLVMTTAYFFLPFLCLSLYGLI 270
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 VEIPAPQDYWGCVFAICIFLFS-----FTIPVLIISVCYSIM 235
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 GRELWSSRRPLGPAASGRERGHOTKRVLLVVLAFIICMLPEHVGRITTYINTEDSRMM 330
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 IKRL-RGVRLLSG--SREKDRNLRIITRLVAVVAVFGCTEPVQV---FVLVQSLGVQ 288
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 331 YFSQYFNIVALQLF-----YLSASINPIILYNLISKYYRAAFKILLARKSRPRGFHRS-- 383
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 PGSE--TAVAILRFCTALGVVNSCLNPIIYAFIDENFKACFRKCCCA-----SALHREMQ 341
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 -RDTAGEVAGDTG 395
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 VSDRVRTVAKDVG 354
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: March 16, 2003, 15:05:27  
Job time : 20 secs

GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:04:24 ; Search time 14 Seconds

(without alignments)  
1356.424 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156  
Sequence: 1 MGSPNMGSDGPGAGARPPWP.....DTGGDTVGYTTSANTYKTMG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 segs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	430	19.9	419	10	US-09-804-551B-26
2	369.5	17.1	362	9	US-09-992-331-13
3	358	16.6	363	9	US-09-992-331-14
4	356.5	16.5	418	9	US-09-992-331-17
5	355.5	16.5	370	9	US-09-966-782A-13
6	349.5	16.2	453	9	US-09-970-966-209
7	349.5	16.2	453	10	US-09-825-294-209
8	340	15.8	380	10	US-09-866-871-77
9	340	15.8	380	12	US-10-039-645-77
10	336.5	15.6	376	10	US-09-966-871-76
11	336.5	15.6	376	12	US-10-039-645-76
12	335.5	15.6	367	9	US-09-823-114-23
13	333.5	15.5	367	9	US-10-087-345A-15
14	333.5	15.5	428	9	US-09-992-331-16
15	331	15.4	372	9	US-09-966-782A-16
16	331	15.4	372	9	US-10-212-980-9
17	330.5	15.3	372	9	US-09-966-871-80
18	330.5	15.3	372	12	US-10-039-645-80
19	330.5	15.3	428	9	US-09-992-331-15

20	330	15.3	376	10	US-09-823-114-17	Sequence 17, Appl
21	330	15.3	364	10	US-09-823-114-22	Sequence 22, Appl
22	329.5	15.3	367	10	US-09-170-919-4	Sequence 4, Appl
23	329.5	15.3	370	10	US-09-823-114-19	Sequence 19, Appl
24	328	15.2	380	10	US-09-214-804-6	Sequence 6, Appl
25	326.5	15.1	371	9	US-09-966-782A-15	Sequence 15, Appl
26	326.5	15.1	371	9	US-10-212-980-8	Sequence 8, Appl
27	325.5	15.1	372	9	US-10-112-559A-4	Sequence 4, Appl
28	321.5	14.9	370	10	US-09-823-114-21	Sequence 21, Appl
29	320.5	14.9	368	9	US-10-212-980-2	Sequence 2, Appl
30	319.5	14.8	368	9	US-09-966-782A-14	Sequence 14, Appl
31	319.5	14.8	368	9	US-10-212-980-4	Sequence 4, Appl
32	319.5	14.8	427	9	US-10-007-132-4	Sequence 4, Appl
33	319	14.8	372	9	US-10-112-559A-2	Sequence 2, Appl
34	319	14.8	372	10	US-09-214-804-4	Sequence 4, Appl
35	319	14.8	380	9	US-10-166-359-2	Sequence 2, Appl
36	319	14.8	380	9	US-10-166-113-2	Sequence 2, Appl
37	319	14.8	380	9	US-10-166-357-2	Sequence 2, Appl
38	319	14.8	380	9	US-10-166-372-2	Sequence 2, Appl
39	318.5	14.8	380	10	US-09-149-045-2	Sequence 2, Appl
40	318.5	14.8	387	9	US-09-966-782A-17	Sequence 17, Appl
41	318.5	14.8	387	9	US-10-212-980-10	Sequence 10, Appl
42	317.5	14.7	370	9	US-10-007-132-2	Sequence 2, Appl
43	316	14.7	362	9	US-10-087-345A-17	Sequence 17, Appl
44	315	14.6	383	10	US-09-966-871-87	Sequence 87, Appl
45	315	14.6	383	12	US-10-039-645-87	Sequence 87, Appl

## ALIGNMENTS

RESULT 1  
US-09-804-551B-26  
Sequence 26, Application US/09804551B  
Patent No. US20020056151A1  
GENERAL INFORMATION:  
APPLICANT: Bayer Aktiengesellschaft  
TITLE OF INVENTION: Receptors for peptides from insects  
FILE REFERENCE: Le A 34 394  
CURRENT APPLICATION NUMBER: US/09/804,551B  
PRIORITY FILING DATE: 2001-03-12  
PRIORITY APPLICATION NUMBER: DE 100 13 618.4  
PRIORITY FILING DATE: 2000-03-18  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 26  
LENGTH: 419  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-804-551B-26

Query Match 19.9%; Score 430; DB 10; Length 419;  
Best Local Similarity 27.7%; Pred. No. 1.4e-27;  
Matches 119; Conservative 61; Mismatches 119; Indels 130; Gaps 10;

QY	23	PCDDRRCSPPGALVPTAVCLFLVGVSGNVVYMLIGRYDMRTTNLYGSMNV	82
DB	12	PPRD-----PLATVIVPVVYSLFITGVGNISTCIVIKORSMTATNYYLFS	64
QY	83	SDLLIIL-GHPFDLYRMRSPWFGFLCLSLVGBGCTYATLLMHTALSVERYLAI	141
DB	65	SDFLLLSGVQEVSYMSKPYVFGYICIGRLAETSAANAATVLTITATVERVIAIC	124
QY	142	RPLRRLVLRRLRRLAIVLMAVALSAGPFLVGVGQDPGTSVVGNGTARIASP	201
DB	125	HPFLGQMSKLSAIRIIVLWMAIVTAIPQAQFGLHHSVY-----	169
QY	202	LASSPPLMSRAPPPSPSPETAEAALFSRRCRPSAPQALRWML---WVTAYFF	257
DB	170	-----CCGIVRVIVKHSFQSTFTFF	190
QY	258	L-PFLCSTLYGLIGRLWSSRRPLRGPAASGR-----	291

Db 191 LAPMSIIIVLILGIVHLRSTL-VEGPASVBAROOLKSVSDTILLYRGSGTAMSFNG 249  
 QY 292 -----GHRQKRVLLVVLVLAFLICMLPFPHVGRITTY- 322  
 Db 250 GGSAGTAGIAGMGGAQLSSVVRGLNHVGTFRVLRMLVAVVGFELCMAPFHQRLTAIY 309  
 QY 323 -----NTEDSRMVMYFQYFNIVALQFYLSASINPLVNLISKYRAAFKLLARKSR 376  
 Db 310 APARCAKLRDQHEFVYT-VMTYVSVGLYTLSTCINPLLNIMSHKR-EAFKAVLFGKV 367  
 QY 377 PRGFHRSRD 385  
 Db 368 SKGSLSNRN 376

RESULT 2  
 US-09-992-331-13  
 ; Sequence 13, Application US/09992331  
 ; Publication No. US20030022186A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FEDER, JOHN N.  
 ; APPLICANT: MINTIER, GABE  
 ; APPLICANT: RAMANATHAN, CHANDRA S.  
 ; APPLICANT: HAWKEN, DONALD R.  
 ; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,  
 ; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA  
 ; TITLE OF INVENTION: CELLS  
 ; FILE REFERENCE: D0048NP  
 ; CURRENT APPLICATION NUMBER: US/09/992,331  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/308,540  
 ; PRIOR FILING DATE: 2001-07-27  
 ; PRIOR APPLICATION NUMBER: 60/261,782  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/248,483  
 ; PRIOR FILING DATE: 2000-11-14  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 13  
 ; LENGTH: 362  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-09-992-331-13

Query Match 17.1%; Score 369.5; DB 9; Length 362;  
 Best Local Similarity 27.8%; Pred. No. 1e-22;  
 Matches 112; Conservative 58; Mismatches 148; Indels 85; Gaps 13;

QY 5 WNGSDGEGAREPPWALPCCDERRCSPPLGA---LVPTAVCLCLFVVGSGNVVTVM 61  
 Db 12 WNSAASGSHN--WSLVDPVS-----PMGARAVLPVLYLVC--TVGSGNTLVLY 60  
 QY 62 LIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLMRSRPMVFGPLLCLRLSLYVGGC 121  
 Db 61 VVLRARAKKTVTNVYIILNLAADVLFMLGLPF-LATQNAVSWPFGSFLCRLVMTLDGIN 119  
 QY 122 TYATLHMTALSVRRYLAICRPLRARVLTERRRVALIIVLAVALLSAGPFLFVGVGO 181  
 Db 120 QFTSIFCLMWSVDRYLAVVHPLRSARMRPRVAKASAAVWFSLMSLPLLVADVOE 179  
 QY 182 DPGISVVGNGTARIIASSPLASPPPLWLSRAPPPSPGCPETAEMALFSRECRSPAQ 241  
 Db 180 G-----WGTCNLS-----W-----PEPVG 193  
 QY 242 LGALRVMLWTTAVFPLPCLSLIYGLI-----GRELMSRRPLRGPAASGERGHR 294  
 Db 194 LMGAAFTITYSVLGFGLVLCCLYLLIVKKAAGKRVGSSR-----RRRSEP 243  
 QY 295 QTKRVLLVVLVLAFLICMLPFPHVGRITTYIN-TEDSRMVMYFQYFNIVALQFYLSASINPI 353  
 Db 244 KVTIMVVVVVLFVVGCMPLFFIVININLAFTLPEEPTSGALYFFVVLVS--YANSCANPL 301  
 QY 354 LYNLISKYRAAFKLLARKSRPRGFHRSRDTAGVADDTG 396

Db 302 LYGFSLDNFROSFRKALCLR-----RGYGVEDADAIIEPRDKSG 340

RESULT 3  
 US-09-992-331-14  
 ; Sequence 14, Application US/09992331  
 ; Publication No. US20030022186A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FEDER, JOHN N.  
 ; APPLICANT: MINTIER, GABE  
 ; APPLICANT: RAMANATHAN, CHANDRA S.  
 ; APPLICANT: HAWKEN, DONALD R.  
 ; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,  
 ; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA  
 ; TITLE OF INVENTION: CELLS  
 ; FILE REFERENCE: D0048NP  
 ; CURRENT APPLICATION NUMBER: US/09/992,331  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/308,540  
 ; PRIOR FILING DATE: 2001-07-27  
 ; PRIOR APPLICATION NUMBER: 60/261,782  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/248,483  
 ; PRIOR FILING DATE: 2000-11-14  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 363  
 ; TYPE: PRT  
 ; ORGANISM: Rattus norvegicus  
 ; US-09-992-331-14

Query Match 16.6%; Score 358; DB 9; Length 363;  
 Best Local Similarity 27.3%; Pred. No. 8.8e-22;  
 Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

QY 5 WNGSDGEGAREPPWALPCCDERRCSPPLGA---LVPTAVCLCLFVVGSGNVVTVM 61  
 Db 12 WNSAASGSHN--WSLVG-----SASPMGARAVLPVLYLVC--TVGSGNTLVLY 60  
 QY 62 LIGRYDMRTTNLYLGSMAVSDLLILGLPDLVRLMRSRPMVFGPLLCLRLSLYVGGC 121  
 Db 61 VVLRARAKKTVTNVYIILNLAADVLFMLGLPFLATQNAVSWPFGSFLCRLVMTLDGIN 120  
 QY 122 TYATLHMTALSVRRYLAICRPLRARVLTERRRVALIIVLAVALLSAGPFLFVGVGO 181  
 Db 121 QFTSIFCLMWSVDRYLAVVHPLRSARMRPRVAKASAAVWFSLMSLPLLVADVOE 180  
 QY 182 DPGISVVGNGTARIIASSPLASPPPLWLSRAPPPSPGCPETAEMALFSRECRSPAQ 241  
 Db 181 G-----WGTCNLS-----W-----PEPVG 194  
 QY 242 LGALRVMLWTTAVFPLPCLSLIYGLI-----GRELMSRRPLRGPAASGERGHR 294  
 Db 195 LMGAAFTITYSVLGFGLVLCCLYLLIVKKAAGKRVGSSR-----RRRSEP 244  
 QY 295 QTKRVLLVVLVLAFLICMLPFPHVGRITTYIN-TEDSRMVMYFQYFNIVALQFYLSASINPI 353  
 Db 245 KVTIMVVVVVLFVVGCMPLFFIVININLAFTLPEEPTSGALYFFVVLVS--YANSCANPL 302  
 QY 354 LYNLISKYRAAFKLLARKSRPRGFHRSRDTAGVADDTG 396  
 Db 303 LYGFSLDNFROSFRKALCLR-----RGYGVEDADAIIEPRDKSG 341

RESULT 4  
 US-09-992-331-17  
 ; Sequence 17, Application US/09992331  
 ; Publication No. US20030022186A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FEDER, JOHN N.  
 ; APPLICANT: MINTIER, GABE

```

; APPLICANT: RAMANATHAN, CHANDRA S.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: D0048ND
; CURRENT APPLICATION NUMBER: US/09/992,331
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/308,540
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/261,782
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/248,483
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-331-17

```

Query Match 16.5%; Score 356.5; DB 9; Length 418;

Best Local Similarity 26.9%; Pred. No. 1.4e-21; Matches 100; Conservative 63; Mismatches 150; Indels 59; Gaps 9;

```

QY 11 PEGAREPPWPLPCDERRCSPPL-----GALVPTAVCLCFVVGSGNVVTMLIGR 65
DB 15 PENASAMPDPATLGNVAGSPSPAGLAVSGVLIPLVYLVC--VYGLGNSLVIYVLR 71
QY 66 YRDMRTTNLYGMAVSDLLLLGLPDLYLKMRSPVFGGLCRSLSYVEGCTYAT 125
DB 72 HTASBSVTNVLNLADLDELFLGLP-LAQNALSYWPFSLMCRLLWADVGINQFTS 130
QY 126 LHMHTALSEVRYLACRPLARAVLTVRRRRLAIAMAVALLSAGPFLVGVGDDPGI 185
DB 131 IFCLTVMSVDRLAVVHPTRSARWTAAPARTVSAVAWVAWVLPVVYFSGV----- 184
QY 186 SVFGLNGTANIASPPLASPPPLWLSRAPSPSPGPTAEAAALFSRECRSPPAQLGAL 245
DB 185 -----PRGMSTCHNQ-----WPEPALAMRA 204
QY 246 RVMLWTTAYFELPLCLSLYGLIGREIMSSRPLRGPAAGRGHGRQTKRVLLVYL 305
DB 205 GFITYTALGFEPPLVLCYLLIVKVSAGRRVWASCCORRRSERRVTMVAVA 264
QY 306 AFILCMLPFVHGRITLYNTE-DSRMVFSQYFNIVALQLFYLSASINPILYMLISKYRA 364
DB 265 LFLVCMWPFYVLTINIVVCPLEBPAPFGLYFLVALLP--YANSCANPILYGLSTYRFX 321
QY 365 AAFKLLARKSR 376
DB 322 QGFRVLLRPSR 333

```

## RESULT 5

```

US-09-966-782A-13
; Sequence 13, Application US/09966782A
; Publication No. US2003002183A1
; GENERAL INFORMATION:
; APPLICANT: Battaglin, P.
; APPLICANT: Feder, J. N.
; APPLICANT: Mintier, G.
; APPLICANT: Ramanathan, C. S.
; APPLICANT: Westphal, R.
; APPLICANT: Hawken, D. R.
; APPLICANT: Cacace, A.
; APPLICANT: Barber, L.
; APPLICANT: Kornacker, M. G.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY7,
; FILE REFERENCE: D0044NP
; CURRENT APPLICATION NUMBER: US/09/966,782A

```

```

; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,731
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/268,560
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/315,423
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 370
; TYPE: PRT
; ORGANISM: RAT
US-09-966-782A-13

```

Query Match 16.5%; Score 355.5; DB 9; Length 370;

Best Local Similarity 29.9%; Pred. No. 1.4e-21; Matches 115; Conservative 65; Mismatches 133; Indels 71; Gaps 15;

```

QY 31 SPPLGALVPTAVCLCFVVGSGN--VVTVM-----IGRYDMRTTNLYGMAVSDL 85
DB 11 SPGSVGA-VAVPVI FALIFLGMVGNGLVLAVLQPGPSAMQERSTTDLFILNLAVADL 69
QY 86 -LILGLPFD--LYRLMRSPWFGPLCR--LSLYVEGCTYATLHMALSVERYLA 139
DB 70 CFLICVFPQAAIYTL--DAMLFGARVCKVHLIYD--TWYASSFTLAAVSDRYLA 123
QY 140 ICRPLRARVTVRRRRLAIAMAVALLSAGPFLVGVGDDPGISVVPGLNGTARLAS 199
DB 124 VRHPLRRLALTRPNMABANGVWMLAALFSAPYL-----SYGTVRGA 168
QY 200 SPLASPPPLWLSRAPSPSPGPTAEAAALFSRECRSPPAQLGALRVMLWTTAYFPLP 259
DB 169 LELC--VPAM-----EDARRALDVATF--AAGYLLP 196
QY 260 FLCISITLYGIGREIMSSRPLRGPAAGRGHGRQTKRVLLVYLATICMLPHVGR 319
DB 197 VAVVSLAYGRTLCFLMAAVGPAAGAAAEARRRATGRAGRAMLVVVVFGISMLPHV-- 253
QY 320 IYINTEDSR--MMVFSQYFNIVALQLFYLSASINPILYMLISKYRAAFKLLARKSRP 377
DB 254 IHLMAERGAFLTPASFFERTIHACLAYSNSLPLVYLSASHFRRAFRLLPCGRRRH 313
QY 378 RGFHRSRDTAGEVAGDGTGVY 401
DB 314 RHHRRAHRLRVQPASSG--PAGY 336

```

## RESULT 6

```

US-09-970-966-209
; Sequence 209, Application US/09970966
; Patent No. US2002017368A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-966-209

```

Query Match 16.2%; Score 349.5; DB 9; Length 453;

Best Local Similarity 28.5%; Pred. No. 5.7e-21; Matches 111; Conservative 64; Mismatches 170; Indels 45; Gaps 13;

[illegible]

RESULT 7  
US-09-825-294-209  
; Sequence 209, Application US/09825294  
; Patent No. US20020004491A1

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121,484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 453
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-825-294-209

```

Query Match	16.2%	Score 349.5;	DB 10;	Length 453;
Best Local Similarity	28.5%;	Pred. No. 5.7e-21;		
Matches 111;	Conservative 64;	Mismatches 170;	Indels 45;	Gaps 135;

```

QY      20 PALPDCDERCS-----PFLGALVPLV--AVCLLEVVAGSAGVVV---TLIG 60
Db      4 PSLPGSD---CGQIIDHSHVPEEVAATMIKILILVLYIIFVMGLIGSATIRVTVQLOK 60

QY      66 YEDMRITTNLYIGSNAVSPLLI-LIGLPDLYR-LMRSPVWFGELCRSLVYGECTY 123
Db      61 KGYLOKEVDTHWVSLACSDIILVFLIGMPEFISIIINBPLTSSYTLSCKHTFEEACSY 120

QY      124 ATLLHMTALSVERYLAICRPLARVLYTRRRRYALIAVLMAVALLSAGPFLFLVGQDP 183
Db      121 ATLLHVLITLSFERVYLAICRPFYKAVSGCQKLLIGVWTSALVALPLFLFAGTXY-P 179

QY      184 GISVVPGLNGTARIASPLASPPMLSRAPDPPSGPGETAAEALFSSRCRPSPAQ-- 241
Db      180 LVN-VPSHHGLGNRSTSHNHQ-----PETSNNMISITNLSSRWTVQSS 223

QY      242 -IGALRVMLWVTYAFLEFLCLSLIYGLIGELMSSRP--LRGPAASGEBRHQRTKR 299

```

Db	224	TEGAFVYVLVLVLSVAMCMNMQVLMKSQKSLAGCTTRPQLRKGBSEBSRPARQTI	283
Qy	299	VLLVAVLAFILCWLPEFHVGRITV-NTEDSRMMYFSQYNIALQ-LFLYLSASINPL	354
Db	284	FLRLIVVLAVCMWPNQIRIRIMAAKPKHMTSRYSFRAYMILLPFSETFYLSVINPL	343
Qy	355	YNLSIKKYRAAPAKLLILARKSRPRGFRSR	384
Db	344	YTVSSQCFRRVFVQVLLCRSLSLQAHNEKR	373

RESULT 8  
US-09-966-871-77  
Communication INS/09966871

Patent No. US20020127539A1  
 GENERAL INFORMATION:  
 APPLICANT: Koplin, Alan S.  
 TITLE OF INVENTION: Assays for Identifying Receptors Having  
 TITLE OF INVENTION: Alterations in Signaling  
 FILE REFERENCE: 00398/512002  
 CURRENT APPLICATION NUMBER: US/09/366, 871  
 CURRENT FILING DATE: 2001-09-28  
 PRIOR APPLICATION NUMBER: US 60/236,302  
 PRIOR FILING DATE: 2000-09-28  
 PRIOR APPLICATION NUMBER: US 60/288,644  
 PRIOR FILING DATE: 2001-05-03  
 NUMBER OF SEQ ID NOS: 87  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 77  
 LENGTH: 380  
 TYPE: prt  
 ORGANISM: Homo sapiens  
 US-09-366-871-77

Query Match	15.8%	Score	340	DB	10	Length	380
Best Local Similarity	25.5%	Pred. No.	2.7e-20				
Matches	103	Conservative	75	Mismatches	146	Indels	80
						Gaps	13

```

QY      6  NGSODPEARREP---WALPPEOBERCSPEPLGALPVPATVCLCEPVGVSNGVTVML 62
Db      39  NGSVSESDQOLEPAHISPAIP-----VITAVYSVTVVGLVGNLSLMFV 83
QY      63  IGRVDEMRRTNNILYGSMAVSDLLILGLPELYRLWRSRPWEGPLLCRLSYVEGCT 122
Db      84  IIRYRKMTAINIYIFNLALADALTVTMTMPQ--SAVYLMNSMPGDVLCXIVISIDYMM 142
QY      123  YATLLHMTALSVERYLAIQPLPRARVLVTRRRVALTAVLMAVALSAGFPLVVGEO 182
Db      143  FTSIFLLTMSSYDRJIAVCHVKAIDFETPLKAKINICIMIIALSVSIGISAIVLGGKVR 202
QY      183  PGISVPGNGINARLASGLASSPPLMSRAPPSPGSPETAEMALFSERCBPAPOL 242
Db      203  EDVDVIE-----CSIQPFDDIYSWMD-----LPMKIC----- 229
QY      243  GALPVLMTWTAYFPLPFLCISILYGLIGRELWSSRBLRGPASGREGHROTKVLLV 302
Db      230  -----VFVAFPIFPLIIIVCTLMILRL-KSVRLTSG--SREKDRNLRRTIKIVL 278
QY      303  VVLAFLICMLPHVGRITVI--NTEDESMYTFQYFNIVALQLFYLASINPLIYNLSK 366
Db      279  VVAVFICWTPIPHILFVLEALGSTSHSTAVLSYF--CALDQYSSINPLVAFLE 335
QY      361  KYRAA-----AFKLLARKRPGRFHRSRDTAGVAG--DTGG 396
Db      336  NFKRCRDFCEPIKMEROST---NVRNIVODPMSMRDVG 375

```

RESULT 9  
US-10-039-645-77  
; Sequence 77, Application US/10039645  
; Patent No. US20020147170A1  
; GENERAL INFORMATION:



APPLICANT: Kopin, Alan S.  
 APPLICANT: Beinborn, Martin  
 TITLE OF INVENTION: Constitutively Active, Hypersensitive,  
 FILE REFERENCE: 00398/510002  
 CURRENT APPLICATION NUMBER: US/10/039,645  
 PRIOR FILING DATE: 2001-10-25  
 PRIOR APPLICATION NUMBER: US 60/243,550  
 NUMBER OF SEQ ID NOS: 87  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 77  
 LENGTH: 380  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-039-645-77

Query Match 15.8%; Score 340; DB 12; Length 380;  
 Best Local Similarity 25.5%; Pred. No. 2.7e-20;  
 Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;

QY 6 NGSDGEGAREPP--WPALPPCERRCSPFLGALVPTAVCLCFVVGSGNVTVML 62  
 DB 39 NGVSGEDQLEPAHISAIP-----VITTAIVSVFVVGVLGNSLVMFV 83  
 QY 63 IGRVDRRTTNLYGSAVSDLLILGLPDLRYLRMSRPVWFGELCRSLVYGEQCT 122  
 DB 84 IIRYTKMTATNIYIFNLADALVTTMPFO-SAVYLMNSWPGVGLCKIVISIDYNNM 142  
 QY 123 YATLHMTALSVRYLACRPLARVLTVRRAVRLANLMLVALLSAPFLFVVEOD 182  
 DB 143 FTSFTLTMVVDRIACHPVKALDFRTPLAKATINICIMTLASSVGISALVIGTKR 202  
 QY 183 PGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPGPTAEAAALFSPRCRPSAQL 242  
 DB 203 EDVVDIE-----CSLQFPDDDEXWMD-----LFMKIC----- 229  
 QY 243 GALKVMLWTATYFPLPCLISILYGLIGRELMSSRRPLRGPAASGRERGRHQRKVLIV 302  
 DB 230 -----VFVFAFVLPVLIIVCYTLMILRL-KSVRLSG--SREKDRNLRTKIVLV 278  
 QY 303 VLAFLICMLPEHGRITVI--NTEDSRMYSOYENIVALOLFSLASINILYLVK 360  
 DB 279 VVAVFIICWITPIHIFILVEALGSTSHSTAVSSYF--CIALGTNSSLNVLVAFIDE 335  
 QY 361 KYRAA-----AFKLLARKSRPRGFHRSRDTAGVAG--DTGG 396  
 DB 336 NFKRCFRDPCFPIKMMEROST---NRVNTVQDPASMRDVG 375

RESULT 10  
 US-09-966-871-76  
 Sequence 76, Application US/09966871  
 Patent No. US20020127539A1  
 GENERAL INFORMATION:  
 APPLICANT: Kopin, Alan S.  
 TITLE OF INVENTION: Assays for Identifying Receptors Having  
 Alterations in Signaling  
 FILE REFERENCE: 00398/512002  
 CURRENT APPLICATION NUMBER: US/09/966,871  
 CURRENT FILING DATE: 2001-09-28  
 PRIOR APPLICATION NUMBER: US 60/236,302  
 PRIOR FILING DATE: 2000-09-28  
 PRIOR APPLICATION NUMBER: US 60/288,644  
 PRIOR FILING DATE: 2001-05-03  
 NUMBER OF SEQ ID NOS: 87  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 76  
 LENGTH: 376  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-966-871-76

Query Match 15.6%; Score 336.5; DB 10; Length 376;  
 Best Local Similarity 24.6%; Pred. No. 5.2e-20;  
 Matches 102; Conservative 76; Mismatches 158; Indels 79; Gaps 12;

QY 1 MGSPWNGSDGEGAREPPWPAALP-----CDERRCSPFLGALVPV- 41  
 DB 1 MESPFGEPTCA--PSACLPPNSSAMFPGMAEPSSNGSAGEDAQLEPAHISPADVE 57  
 QY 42 -TAVCLCFVVGSGNVTVMLIGRYDMRTTNLYGSAVSDLLILGLPDLRYLMR 100  
 DB 58 ITAVSVFVVGVLGNSLVMFVIIRYTKMTATNIYIFNLADALVTTMPFO-STVYL 116  
 QY 101 SRPWFGPLCRSLVYGEQCTVATLHMTALSVRYLACRPLARVLTVRRAVRLA 160  
 DB 117 MNSWPGDVLCIVISIDYNNMFTSIFTLTMSVDRIACHPVKALDFRTPLAKATINI 176  
 QY 161 VLAVALISAGPLFVGVVEODPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPDS 220  
 DB 177 CIMLSSVGISALVIGTGYREDVDIE-----CSLQFPDDDEXWMD----- 219  
 QY 221 GPETAAALFSPRCRPSAQLGALRVMLWTATYFPLPCLISILYGLIGRELMSSRRP 280  
 DB 220 -----LFMKIC-----VFVFAFVLPVLIIVCYTLMILRL-KSVRL 254  
 QY 281 LRGPASGRERGRHQRKVLIVLVVLAFLICMLPEHGRITVI--NTEDSRMYSOYENI 338  
 DB 255 LSG--SREKDRNLRTKIVLVVAVFVCGMTPIHIFILVEALGSTSHSTAVSSYF-- 310  
 QY 339 VALQFLYSINPILYLVNLSKRYRAA-----AFKLLARKSRPRGFHRSRDTA 387  
 DB 311 -CIALGTNSSLNVLVAFIDENFKRCFRDPCFPLKMMEROSTSRVNTVQDPA 364

RESULT 11  
 US-10-039-645-76  
 Sequence 76, Application US/10039645  
 Patent No. US20020147170A1  
 GENERAL INFORMATION:  
 APPLICANT: Kopin, Alan S.  
 TITLE OF INVENTION: Constitutively Active, Hypersensitive,  
 FILE REFERENCE: 00398/510002  
 CURRENT APPLICATION NUMBER: US/10/039,645  
 CURRENT FILING DATE: 2001-10-25  
 PRIOR APPLICATION NUMBER: US 60/243,550  
 PRIOR FILING DATE: 2000-10-26  
 NUMBER OF SEQ ID NOS: 87  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 76  
 LENGTH: 376  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-039-645-76

Query Match 15.6%; Score 336.5; DB 12; Length 376;  
 Best Local Similarity 24.6%; Pred. No. 5.2e-20;  
 Matches 102; Conservative 76; Mismatches 158; Indels 79; Gaps 12;

QY 1 MGSPWNGSDGEGAREPPWPAALP-----CDERRCSPFLGALVPV- 41  
 DB 1 MESPFGEPTCA--PSACLPPNSSAMFPGMAEPSSNGSAGEDAQLEPAHISPADVE 57  
 QY 42 -TAVCLCFVVGSGNVTVMLIGRYDMRTTNLYGSAVSDLLILGLPDLRYLMR 100  
 DB 58 ITAVSVFVVGVLGNSLVMFVIIRYTKMTATNIYIFNLADALVTTMPFO-STVYL 116  
 QY 101 SRPWFGPLCRSLVYGEQCTVATLHMTALSVRYLACRPLARVLTVRRAVRLA 160  
 DB 117 MNSWPGDVLCIVISIDYNNMFTSIFTLTMSVDRIACHPVKALDFRTPLAKATINI 176  
 QY 161 VLAVALISAGPLFVGVVEODPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPDS 220

```

Db 177 CWMSSVGAIVLGGTKREDVDVIE-----CSIQFPDDDDYKMD----- 219
QY 221 GPETAAMALFSGRECPSPAQALRWMLWTTAYFPLPCLSLIYGLIGRELSSRRP 280
Db 220 -----LFFKIC-----VFIFAFVILPVLIIIVCTLMILRL-KSVRL 254
QY 281 LRGPASGERGHRQTKRVLLVVLAFIICWLPFHGRIIYI--NNEDSMMWFSCYFNI 338
Db 285 LSG--SREKDRMLRITRLVLVVAVFVGCWTPVQV--FILTVEALGSIHSTALSIYF-- 310
QY 339 VALQFYLASINPLIYNLISKYRAA-----AFKLLARKSRPGRFHSRDTA 387
Db 311 -CIALGYNSSLNPLIYAFLDENFKCFRDFCPLIMRMERGOSTSRVRLTVQDPA 364

RESULT 12
US-09-823-114-23
; Sequence 23, Application US/09823114
; Patent No. US20020061554A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, CHRISTOPHER J.
; KEITH, DUANE E.
; TITLE OF INVENTION: OPIOID RECEPTOR GENES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & ROESTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/823,114
; FILING DATE: 29-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/148,351
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MORASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20526.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSNROERSMSH
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: group(19, 26, 37)
; OTHER INFORMATION: /note="extracellular Asn residues
; that are consensus sites for N-linked glycosylation"
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-823-114-23

Query Match 15.6%; Score 335.5; DB 10; Length 367;
Best Local Similarity 27.4%; Pred. No. 6.1e-20;
Matches 101; Conservative 68; Mismatches 132; Indels 67; Gaps 11;

QY 34 PLGALVPYAVCLCFVVGSGNVVTVMILGIRYDRMTTNNYLGSAVSDLLILGLPF 93
Db 44 PLGKVTIVGLVAVCGILGNCLVMYVILHHTKTKATNIIYFNLADTLVLTLPF 103

```

```

QY 94 ----DYELMRSPWFGPLLCRLSLVYEGCTYATLLHMTALSVERYIATCPRLARVLV 150
Db 104 QGTDLI-----LGPWFENALCKTVIADYNNMFSTFLTMSYDRVAICHPIRALDVR 159
QY 151 TRRRVRALIVAAVALLSGPFLVGVGEBODPISVPGNGTARLASSPLASSPLML 210
Db 160 TSSKAQAVNVAIVMALASV-----VGVPVAIMSAQVDEDEIEC-----L 198
QY 211 SRAPPPSPSGPETAAMALFSGRECPSPAQALRWMLWTTAYFPLPCLSLIYGLI 270
Db 199 VEIPTPDYWGVPVAFICFLFS-----FIVPLVISVCYSLM 235
QY 271 GRELSSRRPLRGPASGERGHRQTKRVLLVVLAFIICWLPFHGRIIYINTEDSMM 330
Db 236 IRRLLRGVRLLSG--SREKDRMLRITRLVLVVAVFVGCWTPVQV---FVLAQGLGVQ 288
QY 331 YFSQYFNIVALQF-----YLSASINPLIYNLISKYRAAFAKLLARKSR--PGFHS 383
Db 289 PSSR--TAVAILRFTALGVNSCLNPLIYAFLDENFKACRKFCCASALRDRVQSDRV 346
QY 384 RDTAGEVA 391
Db 347 RSIKQDVA 354

RESULT 13
US-10-087-345A-15
; Sequence 15, Application US/10087345A
; Publication No. US20030045696A1
; GENERAL INFORMATION:
; APPLICANT: Ouyang, Chungh
; TITLE OF INVENTION: ORPHANIN FQ RECEPTOR NUCLEIC ACIDS
; FILE REFERENCE: UM-06962
; CURRENT APPLICATION NUMBER: US/10/087,345A
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 15
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-087-345A-15

Query Match 15.5%; Score 333.5; DB 9; Length 367;
Best Local Similarity 26.8%; Pred. No. 8.9e-20;
Matches 100; Conservative 68; Mismatches 132; Indels 73; Gaps 12;

QY 34 PLGALVPYAVCLCFVVGSGNVVTVMILGIRYDRMTTNNYLGSAVSDLLILGLPF 93
Db 44 PLGKVTIVGLVAVCGILGNCLVMYVILHHTKTKATNIIYFNLADTLVLTLPF 103
QY 94 ----DYELMRSPWFGPLLCRLSLVYEGCTYATLLHMTALSVERYIATCPRLARVLV 150
Db 104 QGTDLI-----WPGNALCKTVIADYNNMFSTFLTMSYDRVAICHPIRALDVR 159
QY 151 TRRRVRALIVAAVALLSGPFLVGVGEBODPISVPGNGTARLASSPLASSPLML 210
Db 160 TSSKAQAVNVAIVMALASV-----VGVPVAIMSAQVDEDEIEC-----L 198
QY 211 SRAPPPSPSGPETAAMALFSGRECPSPAQALRWMLWTTAYFPLPCLSLIYGLI 270
Db 199 VEIPTPDYWGVPVAFICFLFS-----FIVPLVISVCYSLM 235
QY 271 GRELSSRRPLRGPASGERGHRQTKRVLLVVLAFIICWLPFHGRIIYINTEDSMM 330
Db 236 IRRLLRGVRLLSG--SREKDRMLRITRLVLVVAVFVGCWTPVQV---FVLAQGLGVQ 288
QY 331 YFSQYFNIVALQF-----YLSASINPLIYNLISKYRAAFAKLLARKSRPGRFHS-- 383
Db 289 PSSR--TAVAILRFTALGVNSCLNPLIYAFLDENFKACRKFCCA-----SSLHREMO 341
QY 384 -RDTAGEVAGDTG 395

```

Db 342 VSDRVSTAKDVG 354

# RESULT 14

US-09-992-331-16  
Sequence 16, Application US/09992331  
Publication No. US20030022186A1

## GENERAL INFORMATION:

APPLICANT: FEDER, JOHN N.  
APPLICANT: MINTIER, GABE  
APPLICANT: RAMANATHAN, CHANDRA S.

APPLICANT: HAWKEN, DONALD R.

TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBWM718,  
TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA

TITLE OF INVENTION: CELLS

FILE REFERENCE: D0048NP

CURRENT APPLICATION NUMBER: US/09/992,331

CURRENT FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: 60/308,540

PRIOR FILING DATE: 2001-07-27

PRIOR APPLICATION NUMBER: 60/261,782

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: 60/248,483

PRIOR FILING DATE: 2000-11-14

NUMBER OF SEQ ID NOS: 45

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 428

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-992-331-16

Query Match

Best Local Similarity 15.5%; Score 333.5; DB 9; Length 428;  
Matches 93; Conservative 64; Mismatches 129; Indels 67; Gaps 9;

36 GALPVTAVCCLEFVGVSGNVVTMLGRYRDMRTTNLYGSAVSDLLILGLPPL 95

45 GILISLYLVVC--VGLGSLSVLYVLRHSSPSVTSVYTLNALDELMLGPF-L 101

96 YLRMSRPWVFGPLLRLSLVVGEGCTATLLHMTALSVERYLACRPLRARVLTTRRV 155

102 AAGNALSYWPGSLMCRVMAVDGNGTSTFCCLTMSVDRLAVVHPRKSARMTAPYA 161

156 RALIAVMAVALLSAGPPLFLVGVQDPGISVVPGLNGTARIASSPLMSRAP 215

162 RMVSAAVWVASAVVVLVVPVFSGV----- 185

216 PSPPSGPEETAALFSGRER---PSPQGLALRMVMTTAYFFLPCLSLIYGLIGR 272

186 ---PRGMS-----CHMWPBPAAAMRTAFITYTALGFPGPLVLCCLYLLIV 232

273 ELWSSRRPLR-----GPAASGRERGHROTKEVLVAVVLAFLICMLPFHVGRIITYNT 324

233 KYRSTTRVAPSCQWQAPACQRRRSERVRVTMVAVALFLVCMWPFYTLNIVNVVC 292

325 E-DSKMYFSCQYFIVVALQFLYSASINPIILYNTLSKKYRAAFAKLLIARSR 376

293 PLDEEPAFGLVPLVALP--YANSCANPILYGLISYRFK--QGFRRIILRPSR 342

# RESULT 15

US-09-966-782A-16  
Sequence 16, Application US/09966782A  
Publication No. US20030022183A1

## GENERAL INFORMATION:

APPLICANT: Battaglino, P.

APPLICANT: Feder, J. N.

APPLICANT: Minter, G.

APPLICANT: Ramanathan, C. S.

APPLICANT: Westphal, R.

APPLICANT: Hawken, D. R.

APPLICANT: Cacace, A.

APPLICANT: Barber, L.

APPLICANT: Kornacker, M. G.

TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBWM71,

TITLE OF INVENTION: EXPRESSED HIGHLY IN SPINAL CORD

FILE REFERENCE: D0044NP

CURRENT APPLICATION NUMBER: US/09/966,782A

CURRENT FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 60/235,731

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: 60/268,580

PRIOR FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: 60/315,423

PRIOR FILING DATE: 2001-08-28

NUMBER OF SEQ ID NOS: 64

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 372

TYPE: PRT

ORGANISM: RAT

US-09-966-782A-16

Query Match

Best Local Similarity 15.4%; Score 331; DB 9; Length 372;  
Matches 110; Conservative 62; Mismatches 133; Indels 96; Gaps 16;

6 NSDGPGEAREP---PWPALPCDERRCSPPLGALVPTAVCCLEFVGVSGNVVTVM 61

4 SSGQGAENTSQEGSGGQWQ-----PBAVLVPL--PFALLFVGTGVNALVLA 48

62 LIGRYDMRTTNLYGSAVSDLLILGLPPL--LYLRMSRPWVFGPLLRLSLVYG 118

49 VLIRGQAVSTTNLFTINLGVADLCFLLCCVFOATYTL--DDWVFGSLCKAVHFLI 105

119 EGCTVATLLHMTALSVERYLACRPLRARVLTTRRVRLALIVMAVALLSAGPFLVIG 178

106 FLTMHSSFTLAASLDRIYAIRPLHSRELRTPRNALAIIGLWGLALFSGGYLSYR 165

179 VEDPGISVVPGLNGTARIASSPLMSRAPPSGPEETAALFSGRERPS 238

166 OSQLANLVV-----CHPAM-----SAPRR----- 184

239 PAQGLARVMTTAY--FFLPCLSLIYGLIGREIWSRRPLRGPAASGRERGHROT 297

185 -----RAMDLCFVFSYLLPVLVLTATRTLRVLTVDV--TAGSGSGQAKRKVT 235

298 RVLVAVVLAFLICMLPENV-----GRIIYINTEDSRMYFSCQYFIVVALQFLYSASI 350

236 RMIIIVAVFLCWMVPHHALILCWFGRPPL-----TRATY--ALRISHLVYANSCV 287

351 NPILYNTLSKKYRAAFAKLLIARSRPRGPHRSRDTAGEVA 391

288 NPIVVALVSKHFR-KGFRKICAGLKP-----APRRASGRVS 323

Search completed: March 16, 2003, 15:08:53  
Job time: 16 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:13 ; Search time 17 Seconds

(without alignments)  
713,073 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156

Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGGDTGYTETSANVKTWG 412

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	857.5	39.8	366	4	US-09-077-675A-13
2	857	39.7	353	4	US-09-077-675A-3
3	855.5	39.7	364	4	US-09-077-675A-16
4	854	39.6	361	4	US-09-077-675A-8
5	769.5	35.7	302	4	US-09-077-675A-2
6	768.5	35.6	302	4	US-09-077-675A-7
7	668.5	31.0	271	4	US-09-077-675A-12
8	638.5	29.6	289	4	US-09-077-675A-10
9	627	29.1	289	4	US-09-077-675A-5
10	477	22.1	415	4	US-09-545-944-2
11	463	21.5	353	1	US-08-118-270-45
12	463	21.5	353	5	PCT-US93-08528-45
13	407	18.9	410	3	US-08-858-876A-2
14	407	18.9	410	3	US-09-472-880-2
15	401	18.6	416	3	US-08-858-876A-4
16	401	18.6	416	4	US-09-472-880-4
17	391.5	18.2	410	4	US-09-200-090-2
18	382	17.7	391	4	US-09-200-090-2
19	358.5	16.6	319	3	US-08-832-339-2
20	358.5	16.6	319	4	US-09-372-438-2
21	356.5	16.5	418	1	US-07-816-283-10
22	356.5	16.5	418	1	US-08-417-103-10
23	338	15.7	380	4	US-08-168-275A-5
24	338	15.7	380	4	US-09-351-158-5
25	338	15.7	380	4	US-09-113-426-5
26	335.5	15.6	367	3	US-08-676-351-2
27	335.5	15.6	367	4	US-08-405-271A-23

28	335	15.5	352	4	US-09-029-027B-2	Sequence 2, Appl
29	334	15.5	380	4	US-08-765-743-2	Sequence 2, Appl
30	333.5	15.5	367	2	US-08-454-549-2	Sequence 2, Appl
31	333.5	15.5	367	3	US-08-454-552-2	Sequence 2, Appl
32	333.5	15.5	367	3	US-08-147-592A-6	Sequence 6, Appl
33	333.5	15.5	367	3	US-08-889-108-17	Sequence 17, Appl
34	333.5	15.5	367	4	US-08-293-694A-6	Sequence 6, Appl
35	333.5	15.5	367	5	PCT-US94-10358-17	Sequence 17, Appl
36	333	15.4	380	3	US-08-676-351-5	Sequence 5, Appl
37	331	15.4	372	2	US-08-626-685A-8	Sequence 8, Appl
38	331	15.4	372	4	US-08-993-088A-2	Sequence 2, Appl
39	331	15.4	372	4	US-08-993-424B-2	Sequence 2, Appl
40	331	15.4	372	4	US-08-665-034A-4	Sequence 4, Appl
41	330.5	15.3	367	3	US-09-170-331-4	Sequence 4, Appl
42	330.5	15.3	428	1	US-07-816-283-12	Sequence 12, Appl
43	330.5	15.3	428	1	US-08-417-103-12	Sequence 12, Appl
44	330	15.3	376	4	US-08-387-707-17	Sequence 17, Appl
45	330	15.3	376	4	US-08-405-271A-17	Sequence 17, Appl

## ALIGNMENTS

RESULT 1  
US-09-077-675A-13  
Sequence 13, Application US/09077675A  
Patent No. 6242199

## GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pond, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ

COUNTRY: USA  
ZIP: 07065-0900

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: PASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720  
TELEX:

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
LENGTH: 366 amino acids

TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-077-675A-13

Query Match 39.8%; Score 857.5; DB 4; Length 366;  
Best Local Similarity 44.5%; Pred. No. 1,5e-65;

Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

QY 5 WNGSDPEGA-----REPPWALPPCD---ERRCSPFPGALVPTAVCLCFVGVSGN 56  
 DB 2 WNAFSEBPGFNLTLADLDWDASPGNDSLGDLELLQFPAPPLAGVATGVALFVVGIAGN 61

QY 57 VVTWMLIGRYRMRTTNNIYLSMAVSDLLILGLPFDLYLRMSRPWFGPLCRSLY 116  
 DB 62 LLTLMVSRFRLRTTNNIYLSMAFSDLLIFCMLPDLRLMQRPWNGDLCKLPQ 121

QY 117 VGEQCTVATLHMTALSVRYLAICPLRARVLTTRRRVALLIIVMAVALSAGEFLFL 176  
 DB 122 VSECTVATVLTITLSTVRYPFACPLRAKVVTGKRVLVIFVIMAVAFSAGPIFVL 181

QY 177 VGEVDPGIVSVVGLNGTARIASSPLASSPPLMSRAPPPSPSGPETAALFSPRECR 236  
 DB 182 VGVEHE-----NGT-----DP--W-----DTNECR 199

QY 237 PS-PAQGLALVLMVMTAVFPLPLCLILYGLIGRELMSRRPLRGPASGRERGR 294  
 DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCLVLYSLIGRKLRRRGDAVVGASLRDQNHK 259

QY 295 QCRVLLVVLVLAIFICMLPFHVRIIYINTEDS---RMVFSQYFNIALQFLYLSASIN 351  
 DB 260 QIVKMLAVVFAFICMLPFHVGRYLFPSKSFEPGSLIAQISQCNLVSFVLYLSAAIN 319

QY 352 PLYNLSKKYRAAFKLLIARKSRPGRHRSRDTAGEVAGDTGDTGVGTETSAN 407  
 DB 320 PLYNMSKKYRAVAFRLGFEPFSQKSLTKDESSR-----AWTESSIN 365

RESULT 2  
 US-09-077-675A-3  
 ; Sequence 3, Application US/09077675A  
 ; Patent No. 6242199  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pai, Lee-Yuh  
 ; APPLICANT: Feighner, Scott C.  
 ; APPLICANT: Howard, Andrew D.  
 ; APPLICANT: Pong, Sheng-Shung  
 ; APPLICANT: Van Der Ploeg, Leonardus H.T.  
 ; TITLE OF INVENTION: RECEPTOR ASSAY  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merck & Co., Inc.  
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
 ; CITY: Rahway  
 ; STATE: NJ  
 ; COUNTRY: USA  
 ; ZIP: 07065-0900  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/077,675A  
 ; FILING DATE: 3-JUN-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Cocuzzo, Anna L.  
 ; REGISTRATION NUMBER: 42,452  
 ; REFERENCE/DOCKET NUMBER: 19590P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 732-594-1273  
 ; TELEFAX: 732-594-4720  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 353 amino acids

TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-077-675A-3

Query Match 39.7%; Score 857; DB 4; Length 353;  
 Best Local Similarity 46.2%; Pred. No. 1,6e-65;  
 Matches 185; Conservative 53; Mismatches 96; Indels 66; Gaps 9;

QY 19 WPAIPCD---ERRCSPFPGALVPTAVCLCFVGVSGNVVWMLIGRYRDMRTTNL 75  
 DB 8 WDAPEPNDLSVEBLLPLFPPPLLAGVATGVALFVVGIAGNLLTLMVSRFRMRTTNL 67

QY 76 YLSMAVSDLLILGLPFDLYLRMSRPWFGPLCRSLYVGEQCTVATLHMTALSVE 135  
 DB 68 YLSMAFSDLLIFCMLPDLRLMQRPWNGDLCKLPQVSECTVATVLTITALSVE 127

QY 136 RYLAICPLRARVLTTRRRVALLIIVMAVALSAGEFLVGVGEQDPGISVVGNGTA 195  
 DB 128 RYFALICPLRAKVVTGKRVLVIMAVAFSAGPIFVLVGVEHD-----NGT- 177

QY 196 RIASSPLASSPPLMSRAPPPSPSGPETAALFSPRECRPS-PAQGLALVLMVMT 253  
 DB 178 -----DPRD-----TWECATERAVASGLITVMWVSS 205

QY 254 AYFPLPCLILYGLIGRELMSRRPLRGPASG---REGRHQTQKVLVVLVLAIFIC 310  
 DB 206 VFPLPFCCLVLYSLIGRKLW---RRKGEAAVGSIRDQNHKTQVLMVAFVFAFIC 262

QY 311 WLPFHVGRIIY---INTEDSRMTFSQYFNIALQFLYLSASINPLNLSKKYRAAF 367  
 DB 263 WLPFHVGRIYLFPSKSLFEPGSLIAQISQCNLVSFVLYLSAAINPLNLSKKYRAAF 322

QY 368 KLLIARKSRPGRHRSRDTAGEVAGDTGDTGVGTETSAN 407  
 DB 323 KLLGEPFSQKSLTKDESSR-----AWTESSIN 352

RESULT 3  
 US-09-077-675A-16  
 ; Sequence 16, Application US/09077675A  
 ; Patent No. 6242199  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pai, Lee-Yuh  
 ; APPLICANT: Feighner, Scott C.  
 ; APPLICANT: Howard, Andrew D.  
 ; APPLICANT: Pong, Sheng-Shung  
 ; APPLICANT: Van Der Ploeg, Leonardus H.T.  
 ; TITLE OF INVENTION: RECEPTOR ASSAY  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merck & Co., Inc.  
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
 ; CITY: Rahway  
 ; STATE: NJ  
 ; COUNTRY: USA  
 ; ZIP: 07065-0900  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/077,675A  
 ; FILING DATE: 3-JUN-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Cocuzzo, Anna L.  
 ; REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 364 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-16

Query Match 39.7%; Score 855.5; DB 4; Length 364;  
Best Local Similarity 48.0%; Pred. No. 2.2e-65;  
Matches 184; Conservative 52; Mismatches 82; Indels 65; Gaps 11;

QY 5 WNGSDGPEGARER-----FWPALPPCD---ERRCSFPLGALVPTAVCLCFVVGVS 55  
DB 2 WNAAT--PSEEPENVTLDLDWDASFGNDSLPDELLPFPAPLAGVTATCAVLFVVGISG 59

QY 56 NVVTWALIRYRDMRTTNLYLGMAVSDLLILGLPEFDLYRLMSRPWVFGPLLCRLSL 115  
DB 60 NLTMLVVRFRFELRTTNLYLGSMASFDDLFLCMLDLVRLMQYRPMWFGDLCFLQ 119

QY 116 YVGECTVATLLHMTALSYERYLAICRPLRRLVTRRRVRLALVAVALLSAGPELF 175  
DB 120 FVSECTVATVLTITLALSYERYPAICRPLRAKVVTGRVGLVILVAVAFSAGPIFV 179

QY 176 LVGVQDPSISVPGNGRARIASPLASSPMLSRAPPPSPSGETLMAALFSEK 235  
DB 180 LVGVHE-----NGT-----DPRD-----TNEC 197

QY 236 RPS--PAQLGALRVLMTVTAFFLPCLSLYLGLIGRELMSRRPLRGPAASG---RE 290  
DB 198 RATEAVNSGLTVMWVSVFFELPVFCITVYSLGRKLM--RR--KDDAVAGSLRD 253

QY 291 RGHROTQKVLVVLVLAFLICMLPFHVGRILYINTEDS---RMWYFSQYFNIVALQLFYLS 347  
DB 254 QNHKQTVKMLAVVFAFILLCMPLFHVGRYLFKSPFGSLIAQISQYCNLVSVFLFYLS 313

QY 348 ASINPIILNLSKKYRAAFKLL 370  
DB 314 AAINPIILNLSKKYRAVAFKLL 336

RESULT 4  
US-09-077-675A-8  
Sequence 8, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077, 675A

FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-8

Query Match 39.6%; Score 854; DB 4; Length 361;  
Best Local Similarity 45.8%; Pred. No. 2.9e-65;  
Matches 182; Conservative 57; Mismatches 98; Indels 60; Gaps 8;

QY 19 WPALPPCD---ERRCSFPLGALVPTAVCLCFVVGVSANVTWMLGRYRDMRTTNL 75  
DB 16 WDASFGNDLSGDELLIQFPAPLAGVTATCAVLFVVGISAGNLTMLVVRFRFELRTTNL 75

QY 76 YLGMAVSDLLILGLPEFDLYRLMSRPWVFGPLLCRLSLYVGECTVATLLHMTALSYE 135  
DB 76 YLSSMAFSDLLIFLCMLPDLVRLMQYRPMWFGDLCFLQFVSECTVATVLTITLALSYE 135

QY 136 RYLAICRPLRRLVTRRRVRLALVAVALLSAGPELFVGVQDPSISVPGNGR 195  
DB 136 RYPAICRPLRAKVVTGRVGLVILVAVAFSAGPIFVLGVHE-----NGT- 185

QY 196 RIASSPPLSLSRAPPPSPSGETLMAALFSEKPS--PAQLGALRVLMTVTT 253  
DB 186 -----DB-W-----DTNECPTERPAVNSGLTVMWVSS 213

QY 254 AYFPLPCLSLYLGLIGRELMSRRPLRGPAASGERGHROTQKVLVVLVLAFLICMLP 313  
DB 214 IFPLPVFCITLVYSLGRKLMRRRGDAVVGASLRDQNHQTVKMLAVVFAFILLCMPL 273

QY 314 FHVGRILYINTEDS---RMWYFSQYFNIVALQLFYLSASINPIILNLSKKYRAAFKLL 370  
DB 274 FHVGRYLFKSPFGSLIAQISQYCNLVSVFLYLSAINPIILNLSKKYRAVAFKLL 333

QY 371 LARKSRPGRFHRSDRTAGEVAGDTGDPVGTETTSAN 407  
DB 334 GFEPFSQRKSLTKDESSR-----AWTESSIN 360

RESULT 5  
US-09-077-675A-2  
Sequence 2, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA

ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 302 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-2

Query Match 35.7%; Score 769.5; DB 4; Length 302;  
Best Local Similarity 46.3%; Pred. No. 3,8e-58;  
Matches 163; Conservative 48; Mismatches 80; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNTLYLGSMAVSDLLILGLPDLRYLRSPRWFGPLICRLSYGE 119  
DB 1 MLVVSFRFRLRTTNTLYLSSMAFSDLLIFLCMPDLVRLMQYRPNWNGNLCKLFQVSE 60  
QY 120 GCTVATLTHMTALSVERTLACRPLRVLVTRRRVRLIALVMAVALSAGPFLVGV 179  
DB 61 SCTVATVLTITLALSVERTYFALCEPLRAKVVTGKRVKVIYIWAFAFCSAGPIFLVGV 120  
QY 180 BODPGISVPGINGTARIASSPLASPPMLSRAPPSPPSGPETAEAALFSRECRPS- 238  
DB 121 EHD-----NGT-----DPRD-----TNECRATE 138  
QY 239 -PAQGLALVMLVTTAVYFPLPCLSLIYLIGRELMSSRRPLRGPAAG--RRGHR 294  
DB 139 FAVRSGLLTVMVWVSIFFLPFCITVLYSLIGRKLW---RRKRGAAVAGSSLRDQNHK 195  
QY 295 QTRVLLVWVLAFLICMLPEHVGRIIY---INTEDSRMWYFSQYFNIALQIFYLSASIN 351  
DB 196 QTVMAVAVFAFLICMLPEHVGRIYLSKSLPQSVETAIQISQYCNLVSVFLFYLSAIN 255  
QY 352 PLYNLISKYRAAFAKLLIARKSRPRGFHRSRDTAGEVAGDTGDTVGTETSAN 407  
DB 256 PLYNLISKYRAVAFRLIGPEPFSQRLSTLKDESSR-----AMTESSIN 301

RESULT 6  
US-09-077-675A-7  
Sequence 7, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
City: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 302 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-7

Query Match 35.6%; Score 768.5; DB 4; Length 302;  
Best Local Similarity 46.2%; Pred. No. 4.7e-58;  
Matches 163; Conservative 51; Mismatches 82; Indels 57; Gaps 7;

QY 60 VMLIGRYDMRTTNTLYLGSMAVSDLLILGLPDLRYLRSPRWFGPLICRLSYGE 119  
DB 1 MLVVSFRFRLRTTNTLYLSSMAFSDLLIFLCMPDLVRLMQYRPNWNGNLCKLFQVSE 60  
QY 120 GCTVATLTHMTALSVERTLACRPLRVLVTRRRVRLIALVMAVALSAGPFLVGV 179  
DB 61 SCTVATVLTITLALSVERTYFALCEPLRAKVVTGKRVKVIYIWAFAFCSAGPIFLVGV 120  
QY 180 BODPGISVPGINGTARIASSPLASPPMLSRAPPSPPSGPETAEAALFSRECRPS- 238  
DB 121 EHD-----NGT-----DP--W-----DTNECRPTE 138  
QY 239 -PAQGLALVMLVTTAVYFPLPCLSLIYLIGRELMSSRRPLRGPAAGRGHRGRTK 297  
DB 139 FAVRSGLLTVMVWVSIFFLPFCITVLYSLIGRKLWRRRDAVAGSLRDQNHKQTV 198  
QY 298 RVLVWVLAFLICMLPEHVGRIIYINTEDS---RMWYFSQYFNIALQIFYLSASINPIL 354  
DB 199 KMLAVAVFAFLICMLPEHVGRIYLSKSLPQSVETAIQISQYCNLVSVFLFYLSAINPIL 258  
QY 355 YNLISKYRAAFAKLLIARKSRPRGFHRSRDTAGEVAGDTGDTVGTETSAN 407  
DB 259 YNLISKYRAVAFRLIGPEPFSQRLSTLKDESSR-----AMTESSIN 301

RESULT 7  
US-09-077-675A-12  
Sequence 12, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.



TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-12

Query Match 31.0%; Score 668.5; DB 4; Length 271;  
Best Local Similarity 44.7%; Pred. No. 1,4e-49;  
Matches 144; Conservative 44; Mismatches 77; Indels 57; Gaps 7;

QY 91 LPFDLYRLMRSPWVGPLLCRLSLVVGEGCTATLHMTALVERLACRPLARVLY 150  
DB 1 MFLDLVRLMQYPMNFGDLCKLFQVVSCTATVTLTALVERFAICPLARAVV 60  
QY 151 TRRYVALIAVMAVALLSAGPFLVGVGODPGISVPGINGTARIASSPPLML 210  
DB 61 TKGRVGLVFIWMAVAFCSAGPIFVLVGEHE-----NGT-----DP--W- 98  
QY 211 SRAPPSPSPSGETAALAFSRECRPS--PAQGLARVMAVTTAVFPLPCLSLIYG 268  
DB 99 -----DINECRPTEFAVASSGLITVWVSSIFFPLPCLTVLYS 138  
QY 269 LIGRELMSRRPLRGPASGRERHQRKRVLLVVLAFICMLPHVVGRIIYNEDS- 327  
DB 139 LIGRLMRRRDVAVGASLRDQNKQTKMLAVVFAFLCLCPHVGVYLSKSKFEFG 198  
QY 328 --RMVFSQYENIVALQLEYLSASINPILYNTISKKYRAAFKLLARKSRPGRHSRD 385  
DB 199 SLEIAQISQYCNLVFVLFYLSAIPILYNTIMSKKYRAVAVFRLLPPEPSQKSLTKD 258  
QY 386 TAGEVAGDTGDTVGYTETSAN 407  
DB 259 ESSR-----AMTESSIN 270

RESULT 8  
US-09-077-675A-10  
Sequence 10, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh

APPLICANT: Feigner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Peng, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-675A-10

Query Match 29.6%; Score 638.5; DB 4; Length 289;  
Best Local Similarity 43.6%; Pred. No. 5,4e-47;  
Matches 143; Conservative 45; Mismatches 79; Indels 61; Gaps 9;

QY 5 WNGSDGEGA-----REPPPALPPCD---BRGSPFLGALVPTAVCLCFVGVSGN 56  
DB 2 WNAIPSEBPGNLTLDADWDASPGNDSLGDELLQLFPAPLAGVTATCVAFVVGIAGN 61  
QY 57 VTYVMLIGRYADMRTTNLYLSNAVSDLLILGLPFDLYRLMRSPWVGPLLCRLSLY 116  
DB 62 LITMLVSRPRELRTTNLYLSNAFSDLLFLCMLPDLVRLMQYPMNFGDLCKLFQ 121  
QY 117 VGSECTATLHMTALVERLACRPLARVLYTRRRVALIAVMAVALLSAGPFLFL 176  
DB 122 VSECTATVTLTALVERFAICPLARAVVTGKGVKLVFIWMAVAFCSAGPIFVL 181  
QY 177 VGVGODPGISVPGINGTARIASSPPLMLSRAPPSPSPSGETAALAFSRECR 236  
DB 182 VGVGEHE-----NGT-----DP--W-----DINECR 199  
QY 237 PS--PAQGLARVMAVTTAVFPLPCLSLIYGILGRELMSRRPLRGPASGRERGR 294  
DB 200 PTEFAVASSGLITVWVSSIFFPLPCLTVLYSLIGRKLRRRRGDAVAGASLRDQNHK 259  
QY 295 QTKRVL-----LVVULA---FITCMLP 313  
DB 260 QTKVLMGSGQALRLSLAGPILSLCLP 287

RESULT 9  
US-09-077-675A-5

```

Sequence 5, Application US/0907675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feigner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-5

Query Match          29.1%; Score 627; DB 4; Length 289;
Best Local Similarity 44.3%; Pred. No. 5, 1e-46;
Matches 117; Conservative 43; Mismatches 71; Indels 58; Gaps 8;

```

```

RESULT 10
US-09-545-944-2
Sequence 2, Application US/09545944
Patent No. 6461836
GENERAL INFORMATION:
APPLICANT: AMES, ROBERT
APPLICANT: ELSHOURBAGY, NABIL
APPLICANT: MICHALOVICH, DAVID
APPLICANT: SARAU, HENRY
APPLICANT: SHABON, USMAN
APPLICANT: VAMIER, LISA
TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
FILE REFERENCE: GP70657-1
CURRENT APPLICATION NUMBER: US/09/545,944
CURRENT FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: US 09/435,384
PRIOR FILING DATE: 1999-11-05
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-545-944-2

Query Match          22.1%; Score 477; DB 4; Length 415;
Best Local Similarity 30.5%; Pred. No. 4, 8e-33;
Matches 113; Conservative 78; Mismatches 102; Indels 78; Gaps 10;

```

```

: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/118,270
: FILING DATE: 09-SEP-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/943,236
: FILING DATE: 10-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Townsend, Kevin G.
: REGISTRATION NUMBER: 34,033
: REFERENCE/DOCKET NUMBER: MURPHY=2A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 353 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-118-270-45

Query Match      21.5%; Score 463; DB 1; Length 353;
Best Local Similarity 32.3%; Pred. No. 6.2e-32;
Matches 122; Conservative 55; Mismatches 111; Indels 90; Gaps 11;

: 39 VPTAVVCLCFVGVGSGNVVYTMIGR---YRDMRTTNLYLGSMAVSDLLILGLPDL 95
: 2 VLVTAIYALFLVGVGNSVTAFTLARKKSLQSLQSTVHYHLSLSLSDLLILMW--EL 59
: 96 YRLRSRPWFVGFLLCRSLYVGEQCTYATLLHMTLSVERYLAIQRPRLRVVTRRV 155
: 60 YNFMHMHMAFGDAGCRGYFLRDCTYATALNVASLSVERYLAIQHPFAKTLMSRSRT 119
: 156 RALIAVLAVALLSAGPFLFVGVQDPGISVVGNGTARIASSPLASSPPLMLSRAP 215
: 120 KKFISAIWLASALLAIPLMFLTLGLQNRSGDGTGPG---GLVCTPIVDT----- 164
: 216 PSPSPGPTAAALFSRECRSPAQOLGALRMVMTTAYFFL--PFLCLSLVGLIGREL 274
: 165 -----ATVKVVIQVNTFMSFLPMLVISIINTVANKL 197
: 275 WSSRRPLRGPAASGR-----ERGRQTKR-----VLVVVLAFLIICW 311
: 198 TWV---VHQAAEGRGVCTVGTNGLEHSTFMNRIBPGVQALRHGVLVRAVAVIAFVVCW 254
: 312 LPEHVGRIIYINTEDSRMYFS--QYFNIVALQLFYLASINPILYNLISKRYRAAFKL 369
: 255 LPY---LCYISDEQWRTEFLDFYHYFMYLNTALFYSSAINPILYNLVSANFRQVFLST 310
: 370 -----LLARKSRP 377
: 311 LACLFCGWPILIRKKRP 328

RESULT 12
PCT-US93-08528-45
; Sequence 45, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

```

```

: NUMBER OF SEQUENCES: 348
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/08528
: FILING DATE: 09-SEP-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/943,236
: FILING DATE: 10-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Townsend, Kevin G.
: REGISTRATION NUMBER: 34,033
: REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SEQ ID NO: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 353 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: PCT-US93-08528-45

Query Match      21.5%; Score 463; DB 5; Length 353;
Best Local Similarity 32.3%; Pred. No. 6.2e-32;
Matches 122; Conservative 55; Mismatches 111; Indels 90; Gaps 11;

: 39 VPTAVVCLCFVGVGSGNVVYTMIGR---YRDMRTTNLYLGSMAVSDLLILGLPDL 95
: 2 VLVTAIYALFLVGVGNSVTAFTLARKKSLQSLQSTVHYHLSLSLSDLLILMW--EL 59
: 96 YRLRSRPWFVGFLLCRSLYVGEQCTYATLLHMTLSVERYLAIQRPRLRVVTRRV 155
: 60 YNFMHMHMAFGDAGCRGYFLRDCTYATALNVASLSVERYLAIQHPFAKTLMSRSRT 119
: 156 RALIAVLAVALLSAGPFLFVGVQDPGISVVGNGTARIASSPLASSPPLMLSRAP 215
: 120 KKFISAIWLASALLAIPLMFLTLGLQNRSGDGTGPG---GLVCTPIVDT----- 164
: 216 PSPSPGPTAAALFSRECRSPAQOLGALRMVMTTAYFFL--PFLCLSLVGLIGREL 274
: 165 -----ATVKVVIQVNTFMSFLPMLVISIINTVANKL 197
: 275 WSSRRPLRGPAASGR-----ERGRQTKR-----VLVVVLAFLIICW 311
: 198 TWV---VHQAAEGRGVCTVGTNGLEHSTFMNRIBPGVQALRHGVLVRAVAVIAFVVCW 254
: 312 LPEHVGRIIYINTEDSRMYFS--QYFNIVALQLFYLASINPILYNLISKRYRAAFKL 369
: 255 LPY---LCYISDEQWRTEFLDFYHYFMYLNTALFYSSAINPILYNLVSANFRQVFLST 310
: 370 -----LLARKSRP 377
: 311 LACLFCGWPILIRKKRP 328

RESULT 13
US-08-858-876A-2
; Sequence 2, Application US/0858876A
; Patent No. 6022856

```

GENERAL INFORMATION:  
APPLICANT: Daniel CAPUT  
APPLICANT: Pascual CHALON  
APPLICANT: Pascual FERRARA  
APPLICANT: Vita NATALIO  
TITLE OF INVENTION: Type 2 Neurotensin Receptor  
TITLE OF INVENTION: (hnt-R2)  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC  
STREET: 400 Seventh Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858,876A  
FILING DATE: 19-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR 9723204  
FILING DATE: 17-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,049  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-858-876A-2

Query Match 18.9% Score 407; DB 3; Length 410;  
Best Local Similarity 31.2% Pred. No. 4.4e-27;  
Matches 129; Conservative 60; Mismatches 135; Indels 90; Gaps 16;  
QY 17 PMPALP---CDERRCPPLGALVPTAVCLCFVVGSGNVTVMLI-----GR 65  
DB 8 PPRPSSPGLSDARLGVDTRLMAKVLFTALYALIMAGAGNALSVHVLKARAGR 67  
QY 66 YRDMRTTNLYGSMANVDLILL-GLPDLYR-LMRSRPWVGPIRLRLSYVGEQCTY 123  
DB 68 LRH-----HVLSTLALAGLLLLVGVPELVSEFVWFHYPMVFGDLGCRGYVVELCAY 120  
QY 124 ATLHMTALSVRYLAICRPLARVLTTRRRVRLAIIVLMAVALLSAGPFLVGV--- 179  
DB 121 ATVLVAGLSAERCLAVCOPLARSLTPRRTRMLVALSMAASLGALPMAVIMQCKHEL 180  
QY 180 -----EODPGISVPGINGTARIASS-----PLASSPPL-----WLS 211  
DB 181 ETADGEPBPASRCVTL--VSRFALQVFIQVNVLVSVLPALTAFINGVTVSHLLALCS 238  
QY 212 RAPPPSPGPGETAALAFSRECRSPPAQGLARVMTLWTTAVFPLPCLSLYGLIG 271  
DB 239 QVPSTSTP-GSSTPSRLLELSEE-----GLLSFIWKKT-----FIQG--G 276  
QY 272 RELMSRRPLRGPAASGERGHROTGRVLLVVLAFIICWLPFHVGRIIYINTEDSR-- 328  
DB 277 QVSLVRHNDVR-----RIRSLQRSVQVLRALIVMIVICWLPFHARLMTCYVDDAMTD 330  
QY 329 -MMYFSQYFNIVALQLFYLSASINPIYNIISKYRAAAKLLARKSRPRGFH 381  
DB 331 PLVNFYHYFVMTNTLFFYVSSAVTPLLNVAVSSFR-----KLFLKAVSSLCGEH 380

RESULT 14  
US-09-472-880-2

Sequence 2, Application US/09472880  
Patent No. 6274333  
GENERAL INFORMATION:  
APPLICANT: Daniel CAPUT  
APPLICANT: Pascual CHALON  
APPLICANT: Pascual FERRARA  
APPLICANT: Vita NATALIO  
TITLE OF INVENTION: Type 2 Neurotensin Receptor  
TITLE OF INVENTION: (hnt-R2)  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC  
STREET: 400 Seventh Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/472,880  
FILING DATE: 28-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR 9723204  
FILING DATE: 17-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,049  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 410 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-472-880-2

Query Match 18.9% Score 407; DB 4; Length 410;  
Best Local Similarity 31.2% Pred. No. 4.4e-27;  
Matches 129; Conservative 60; Mismatches 135; Indels 90; Gaps 16;  
QY 17 PMPALP---CDERRCPPLGALVPTAVCLCFVVGSGNVTVMLI-----GR 65  
DB 8 PPRPSSPGLSDARLGVDTRLMAKVLFTALYALIMAGAGNALSVHVLKARAGR 67  
QY 66 YRDMRTTNLYGSMANVDLILL-GLPDLYR-LMRSRPWVGPIRLRLSYVGEQCTY 123  
DB 68 LRH-----HVLSTLALAGLLLLVGVPELVSEFVWFHYPMVFGDLGCRGYVVELCAY 120  
QY 124 ATLHMTALSVRYLAICRPLARVLTTRRRVRLAIIVLMAVALLSAGPFLVGV--- 179  
DB 121 ATVLVAGLSAERCLAVCOPLARSLTPRRTRMLVALSMAASLGALPMAVIMQCKHEL 180  
QY 180 -----EODPGISVPGINGTARIASS-----PLASSPPL-----WLS 211  
DB 181 ETADGEPBPASRCVTL--VSRFALQVFIQVNVLVSVLPALTAFINGVTVSHLLALCS 238  
QY 212 RAPPPSPGPGETAALAFSRECRSPPAQGLARVMTLWTTAVFPLPCLSLYGLIG 271  
DB 239 QVPSTSTP-GSSTPSRLLELSEE-----GLLSFIWKKT-----FIQG--G 276  
QY 272 RELMSRRPLRGPAASGERGHROTGRVLLVVLAFIICWLPFHVGRIIYINTEDSR-- 328  
DB 277 QVSLVRHNDVR-----RIRSLQRSVQVLRALIVMIVICWLPFHARLMTCYVDDAMTD 330  
QY 329 -MMYFSQYFNIVALQLFYLSASINPIYNIISKYRAAAKLLARKSRPRGFH 381  
DB 331 PLVNFYHYFVMTNTLFFYVSSAVTPLLNVAVSSFR-----KLFLKAVSSLCGEH 380

## RESULT 15

US-08-858-876A-4

Sequence 4, Application US/08858876A

Patent No. 6022856

GENERAL INFORMATION:

APPLICANT: Daniel CAPUT

APPLICANT: Pascale CHALON

APPLICANT: Pascale FERRARA

APPLICANT: Vito NATALIO

TITLE OF INVENTION: Type 2 Neurotensin Receptor

TITLE OF INVENTION: (hnt-R2)

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman &amp; Stern, PLLC

STREET: 400 Seventh Street

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,876A

FILING DATE: 19-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR 9723204

FILING DATE: 17-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,049

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 416 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-858-876A-4

Query Match 18.6%; Score 401; DB 3; Length 416;

Best Local Similarity 29.4%; Pred. No. 1.5e-26;

Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;

QY 18 PWPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTVMLIGRYRDMRT 71

Db 6 PWPRLPSPSAGLSLEBARLGVDTRLMAKVLFTALYSLIFAGTAGNALSVHVLKARAGR 65

QY 72 -TTNLVYLSMAVSDLLILL-GLPFDLYR-LMRSRPWFVFGPLLGRSLVYEGGCTVATLH 128

Db 66 GLRHYHVLISLALSLALLLVSMFMELYNFMVSHYPVFGDLGCRGYFVRELCAVATVLS 125

QY 129 MTALSVERYLAICRPLRARVLTTRRRVRLALVLAVALLSAGPFLVIGV----- 179

Db 126 VASLSAEKCLAVCQPLRRRLITPRTRRLSLVWVASLGLPMAVIMGQKHEVESADG 185

QY 180 EQDGIISVPELNGTARIA-----SSPLASSPPLMLSRAPPSPSGPETAABA 228

Db 186 EPEPASRVCYTLVSRATLQVFQVNVVSPALPLALTAFL-----NGITVNHLM 234

QY 229 ALFERECRSPAPOLGALVLMVMTTAYFFLPCLSLIYGLIGREIWSRRPLRGPAAG 288

Db 235 ALYS-QVPSASAQVSSIPSRLLESE-----EGLIGFITWRKTLISLGVQASLV 281

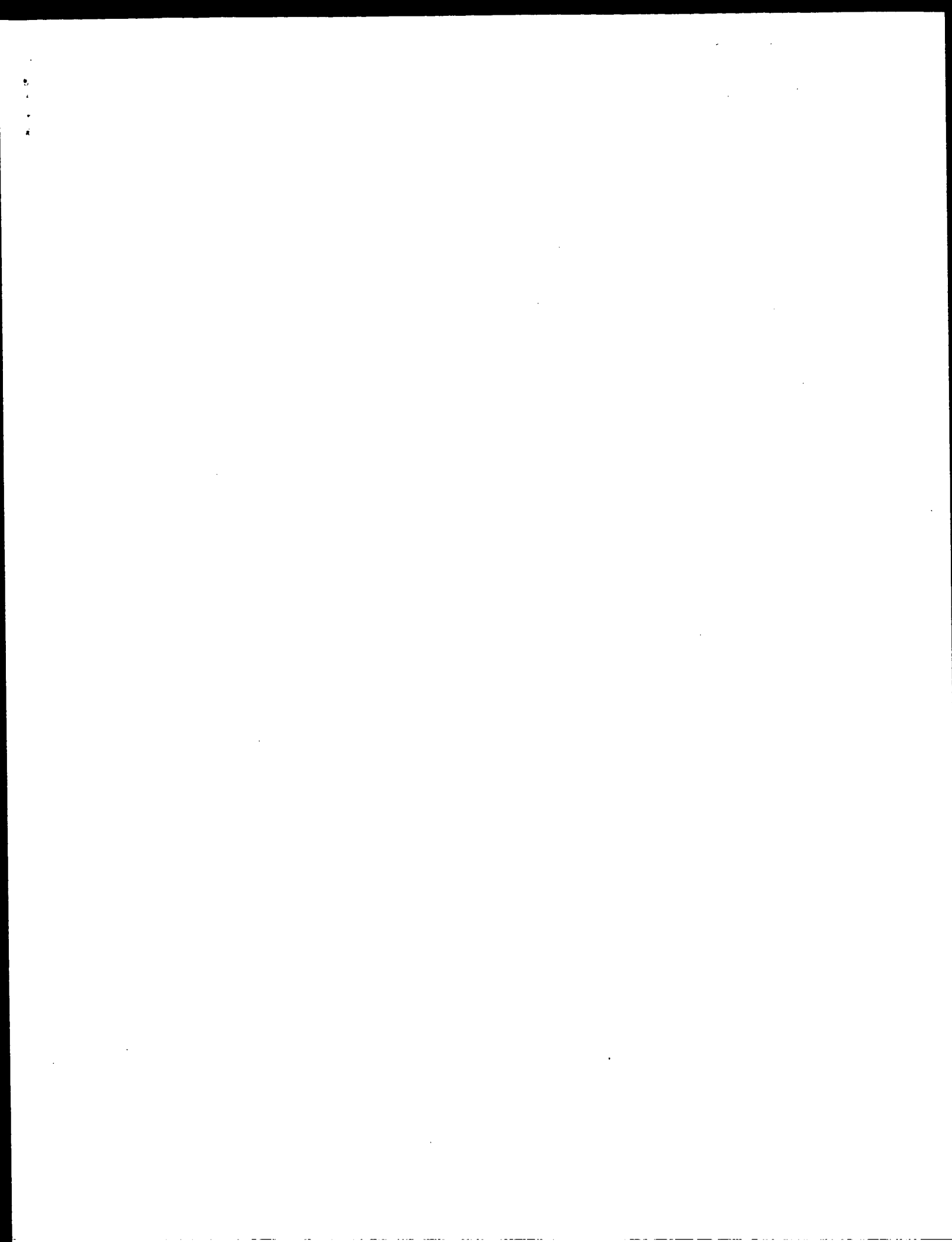
QY 289 PERGHROT-----RVLLVVVLAIIICMLPFGVRIIYINTED-----SRMYFSQYENI 338

Db 282 RHKDAQISRLQHSQAQVLAIVAVVVICMLPFGVRIIYINTED-----SRMYFSQYENI 341

QY 339 VALQFLYLSASINILVNLISKYRAAFAKLLARKSRPRGFHRSRDTAGEVAGD 393

Db 342 VTNLTFYVSSAVTPILYNAVSSFR-----KLFL-----ESLGLGCE 379

Search completed: March 16, 2003, 15:05:54  
Job time : 20 secs



GenCore version 5.1.4\_p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:07 ; Search time 40 Seconds  
(without alignments)  
1372.481 Million cell updates/sec

Title: US-09-876-252-130  
Perfect score: 2156  
Sequence: 1 MGSPNMGSDGPEGAREPPMD.....DTGGDTVGTYETISANKYTMG 412

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```
Database :
_A_Geneseq_101002:*
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:
```

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2156	100.0	412	21	AA802854	Human G protein c
2	2149	99.7	412	21	AA54145	Amino acid sequen
3	2149	99.7	412	22	AA862652	Long form of motil
4	2149	99.7	412	22	AA868478	Amino acid sequen
5	2149	99.7	501	22	AA868622	Human GPR38 varian
6	1741	80.8	400	22	AA868477	Amino acid sequen
7	1575	73.1	386	21	AA54146	Amino acid sequen
8	1575	73.1	386	22	AA862653	Short form of motil
9	1143.5	53.0	271	22	AA868476	Amino acid sequen
10	901	41.8	363	21	AA54147	The pulfer fish mo

11	901	41.8	363	22	AAH68466	Amino acid sequenc
12	864.5	40.1	366	21	AAH90867	Human mutant G pro
13	859.5	39.9	349	21	AAH62923	A canine growth hor
14	857.5	39.8	366	21	AAH70343	Human G protein-cc
15	857.5	39.8	366	21	AAH90435	Human G protein-cc
16	857.5	39.8	366	22	AAH97376	Rat growth hormone
17	857.5	39.8	366	22	AAH62650	Human G-protein cc
18	857	39.7	353	18	AAH19608	Pig growth hormone
19	857	39.7	353	18	AAH19215	Swine growth hormo
20	855.5	39.7	364	21	AAH54565	A mouse growth hor
21	855.5	39.7	364	22	AAH97377	Rat growth hormone
22	854	39.6	361	18	AAH19217	Human growth hormo
23	854	39.6	362	18	AAH19610	Human growth hormo
24	851.5	39.5	364	18	AAH19613	Human growth hormo
25	839.5	38.9	374	18	AAH19220	Rat growth hormone
26	668.5	31.0	271	18	AAH19612	Human growth hormo
27	668.5	31.0	271	18	AAH19219	Human growth hormo
28	635	29.5	289	18	AAH19611	Human growth hormo
29	635	29.5	289	18	AAH19609	Pig growth hormone
30	635	29.5	289	18	AAH19616	Swine growth hormo
31	633.5	29.4	289	18	AAH19218	Human growth hormo
32	498	23.1	418	22	ABB56371	Non-endorogenous hum
33	495	23.0	418	17	AAH98562	Human neurotensin
34	491	22.8	542	22	ABG12316	Novel human diagno
35	490.5	22.8	403	21	AAH90673	Human mutant G pro
36	485.5	22.5	445	22	AAH65823	Human novel cytoKi
37	485.5	22.5	445	22	AAH65866	Human novel cytoKi
38	484.5	22.5	403	21	AAH90638	Human G protein-co
39	484.5	22.5	403	21	AAH46442	Human growth hormo
40	484.5	22.5	403	22	AAH99185	Human FM-3. Homo
41	484.5	22.5	426	22	AAE03928	Human G-protein co
42	481	22.3	415	21	AAH53992	Human neurotensin-
43	481	22.3	415	22	AAH63366	Amino acid sequenc
44	477	22.1	412	22	AAH67805	Amino acid sequenc
45	477	22.1	415	22	AAH67804	Amino acid sequenc

## ALIGNMENTS

RESULT	1
ID	AAB02854
AC	AAB02854 standard; Protein; 412 AA.
XX	
XX	AAB02854;
DT	22-AUG-2000 (first entry)
DE	
XX	Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130
KW	Human; G protein coupled receptor; GPCR; transmembrane receptor; identification; agonist; screening; therapeutic; pharmaceutical; mutant.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200022131-A2.
XX	
PD	20-APR-2000.
XX	
PF	13-OCT-1999; 99MO-US24065.
XX	
PR	13-OCT-1998; 98US-0170496.
PR	12-NOV-1998; 98US-0108029.
PR	20-NOV-1998; 98US-0109213.
PR	27-NOV-1998; 98US-0110060.
PR	16-FEB-1999; 99US-0120416.
PR	26-FEB-1999; 99US-0121852.
PR	12-MAR-1999; 99US-0123944.
PR	12-MAR-1999; 99US-0123945.
PR	12-MAR-1999; 99US-0123946.
PR	12-MAR-1999; 99US-0123948.

PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 28-MAY-1999; 99US-0137567.  
 PR 30-JUN-1999; 99US-0141448.  
 PR 27-AUG-1999; 99US-0151114.  
 PR 03-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156655.  
 PR 29-SEP-1999; 99US-0156634.

(AREN-) ARENA PHARM INC.

Behan DP, Lehmann-Brulinema K, Chalmers DT, Chen R, Dang HT,  
 Gore M, Law CM, Lin I, Lowitz K, White C,

WPI: 2000-317986/27.  
 N-PSDB; AAA46116.

Non-endogenous, human G protein-coupled receptors for screening  
 receptor, inverse or partial agonists useful as therapeutic agents -  
 Example 2; Page 168-169; 187pp; English.

The present invention describes transmembrane receptors, preferably  
 human G protein coupled receptors (GPCR), for which the endogenous  
 ligand is unknown (orphan GPCR receptors). More specifically the present  
 invention relates to non-endogenous, constitutively activated versions  
 of a human GPCR. These non-endogenous human GPCRs can be useful for  
 the direct identification of candidate compounds as receptors agonists,  
 inverse agonists or partial agonists for use as pharmaceutical agents,  
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 the exemplification of the present invention.

SO Sequence 412 AA;

Query Match 100.0%; Score 2156; DB 21; Length 412;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-195;  
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPWPALPCDERRCSPPILGALVPTAVTAVCLCFVVGSGNVTV 60  
 DB 1 MGSPWNGSDGEGARBPWPALPCDERRCSPPILGALVPTAVTAVCLCFVVGSGNVTV 60  
 QY 61 MLIGRYDMRTTNTLYIGSMAVSDLLILGLPFDLYRLMSRPWFVGPILCRSLYVGBG 120  
 DB 61 MLIGRYDMRTTNTLYIGSMAVSDLLILGLPFDLYRLMSRPWFVGPILCRSLYVGBG 120  
 QY 121 CTYATLHMTALSYERYIAICRPLARVLYRRRRALIAVIMANALLSAGFLVNGVE 180  
 DB 121 CTYATLHMTALSYERYIAICRPLARVLYRRRRALIAVIMANALLSAGFLVNGVE 180  
 QY 181 ODPISVVGANGTARIASPPILMSRAPPPSPSGETAAALFSGRECRPSPA 240  
 DB 181 ODPISVVGANGTARIASPPILMSRAPPPSPSGETAAALFSGRECRPSPA 240  
 QY 241 QLGALRWLVMTTAVFPLPCLSLIYGLIGRELWSSRRPLGPAASGERGHRQTKRYL 300  
 DB 241 QLGALRWLVMTTAVFPLPCLSLIYGLIGRELWSSRRPLGPAASGERGHRQTKRYL 300  
 QY 301 LVVTLAFICMLPRVGGIITYINTEDSRMVRSOYENTVALOLFYLASINPLVNLISK 360  
 DB 301 LVVTLAFICMLPRVGGIITYINTEDSRMVRSOYENTVALOLFYLASINPLVNLISK 360  
 QY 361 KYRAAFAFLILARKSRPRGFRHSRDTAGVADGTGDTVGYTETSANVXTMG 412  
 DB 361 KYRAAFAFLILARKSRPRGFRHSRDTAGVADGTGDTVGYTETSANVXTMG 412

RESULT 2  
 AAY54145  
 ID AAY54145 standard; Protein; 412 AA.  
 AC AAY54145;  
 XX  
 DT 27-MAR-2000 (first entry)  
 DE

Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;  
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;  
 KW functional defect; neurological disorder; scleroderma; colonoscopy;  
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;  
 KW infection; stress-related motility disorder; psychogenic disorder;  
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;  
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;  
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;  
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;  
 KW endoscopy; duodenal intubation.

XX Homo sapiens.

XX WO9964436-A1.

XX 16-DEC-1999.

XX 08-JUN-1999; 99WO-US12773.

XX 12-JUN-1998; 98US-0089098.

XX (MERI) MERCK & CO INC.

XX Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;  
 PI Pong S, Smith RG;

XX WPI: 2000-105868/09.

XX N-PSDB; AA245403.

PT Novel receptor protein for screening compounds used in treating  
 irritable bowel syndrome, constipation and other gastric conditions -  
 Claim 3; Fig 3; 44pp; English.

XX The present sequence represents splice variant MTL-R1A of the motilin  
 CC receptor. The gene encodes a G-protein coupled receptor, and is  
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,  
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a  
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated  
 CC five transmembrane domain. The MTL-R1 proteins are used to identify  
 CC agonists and antagonists which can be used for treating gastric motility  
 CC disorders, functional defects, disorders secondary to neurological  
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced  
 CC dysmotility, diabetes, infections, stress-related motility disorders,  
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,  
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal  
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable  
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and  
 CC diarrhoea. They can also be used in the preparation for colonoscopy,  
 CC endoscopy and duodenal intubation.

XX Sequence 412 AA;

Query Match 99.7%; Score 2149; DB 21; Length 412;  
 Best Local Similarity 99.8%; Pred. No. 2.2e-194;  
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPWPALPCDERRCSPPILGALVPTAVTAVCLCFVVGSGNVTV 60  
 DB 1 MGSPWNGSDGEGARBPWPALPCDERRCSPPILGALVPTAVTAVCLCFVVGSGNVTV 60  
 QY 61 MLIGRYDMRTTNTLYIGSMAVSDLLILGLPFDLYRLMSRPWFVGPILCRSLYVGBG 120  
 DB 61 MLIGRYDMRTTNTLYIGSMAVSDLLILGLPFDLYRLMSRPWFVGPILCRSLYVGBG 120



CC hormone secretion in a mammal having a disease associated with abnormal

**XX**



[illegible]

Qy	301	LVVLAFLICMLPFHVGRIITYINTEDSSMMVFESQYENFVALQLEFLSLSNIPILYNIISK	360
Db	289	LVVLAFLICMLPFHVGRIITYINTEDSSMMVFESQYENFVALQLEFLSLSNIPILYNIISK	348
Qy	361	KYRAAARFLLRKSRPRGFHRSRDTAGEVADTGDTGVGTETGSANYKT	410
Db	349	KYRAAARFLLRKSRPRGFHRSRDTAGEVADTGDTGVGTETGSANYKT	398
RESULT 7			
ID	AAV54146		
XX	AAV54146 standard; Protein, 386 AA.		
XX	AAV54146;		
XX	27-MAR-2000 (first entry)		
DE	Amino acid sequence of the motilin receptor splice variant MTL-R1B.		
XX	Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;		
KW	spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;		
KW	functional defect; neurological disorder; scleroderma; colonoscopy;		
KW	paraneoplastic syndrome; radiation induced dysmotility; diabetes;		
KW	infection; stress-related motility disorder; psychogenic disorder;		
KW	gastroenteritis; gastro-oesophageal reflux disease; constipation;		
KW	chronic idiopathic pseudo obstruction; acute faecal impaction;		
KW	postoperative ileus; gallstones; infantile colic; diarrhoea;		
KW	irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;		
XX	endoscopy; duodenal intubation.		
OS	Homo sapiens.		
XX			
PN	WO9964436-A1.		
XX			
PD	16-DEC-1999.		
XX			
PF	08-JUN-1999; 99WO-US12773.		
XX			
PR	12-JUN-1998; 98US-0089098.		
XX			
PA	(MERI ) MERCK & CO INC.		
XX			
PI	Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;		
PI	Pong S, Smith RG;		
XX			
DR	WPI; 2000-105868/09.		
XX			
DR	N-PSDB; AAZ45404.		
XX			
PT	Novel receptor protein for screening compounds used in treating		
PT	irritable bowel syndrome, constipation and other gastric conditions		
XX			
PS	Claim 5; Fig 5; 44pp; English.		
XX			
CC	The present sequence represents splice variant MTL-R1B of the motilin		
CC	receptor. The gene encodes a G-protein coupled receptor, and is		
CC	designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,		
CC	MTL-R1A (see AAV54145) and MTL-R1B (see AAV54146). MTL-R1A is a		
CC	functional seven transmembrane domain form, and MTL-R1B is a truncated		
CC	five transmembrane domain. The MTL-R1 proteins are used to identify		
CC	agonists and antagonists which can be used for treating gastric motility		
CC	disorders, functional defects, disorders secondary to neurological		
CC	disorders e.g. scleroderma, paraneoplastic syndromes radiation induced		
CC	dysmotility, diabetes, infections, stress-related motility disorders,		
CC	psychogenic disorders, gastroparasis, gastro-oesophageal reflux disease,		
CC	constipation, chronic idiopathic pseudo obstruction, acute faecal		
CC	impaction, postoperative ileus, gallstones, infantile colic, irritable		
CC	bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and		
CC	diarrhoea. They can also be used in the preparation for colonoscopy,		
CC	endoscopy and duodenal intubation.		
XX			
SO	Sequence 386 AA;		

Query Match 73.1%; Score 1575; DB 21; Length 386;

Best Local Similarity 99.7%; Pred. No. 3.2e-140;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60  
Db 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60  
QY 61 MLIGRYDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLCRSLVYGE 120  
Db 61 MLIGRYDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLCRSLVYGE 120  
QY 121 CTYATLHMTALSVERYLAICRPLRARVLTTRRRVRLAVLMAVALLSAGPFLVGE 180  
Db 121 CTYATLHMTALSVERYLAICRPLRARVLTTRRRVRLAVLMAVALLSAGPFLVGE 180  
QY 181 QDPGISVVGNGTARIASSPLASPPPLMSRAPPPSPGPEETAALFSECRPSPA 240  
Db 181 QDPGISVVGNGTARIASSPLASPPPLMSRAPPPSPGPEETAALFSECRPSPA 240  
QY 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGHQTVRL 300  
Db 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGHQTVRL 300

RESULT 8  
AAB62653  
ID AAB62653 standard; Protein; 386 AA.

XX AAB62653;

DT 23-JUL-2001 (first entry)

DE Short form of motilin receptor, GPR-38B isoform.

XX zslg33; signal transduction; hormone; enzyme; neural development;  
XX gastric contractility; nutrient uptake; digestive; pancreatic; human;  
XX insulin-like growth factor-II; growth hormone; bone; gastrointestinal;  
XX glucose; osteopathic; anorectic; vulnerability; immunomodulator; GHS-R;  
XX G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.

XX Homo sapiens.

XX WO200138355-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US32074.

XX 22-NOV-1999; 99US-0166765.

XX (ZYMO) ZYMOGENETICS INC.

XX Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

XX WPI; 2001-355879/37.

XX N-PSDB; AAF83684.

XX Forming reversible peptide receptor complex for purifying cell and  
XX peptides, stimulating signal transduction and modulating hormone  
XX secretion, involves contacting a receptor with zslg33 polypeptide  
XX  
XX Disclosure; Page 106-109; 111pp; English.

XX The invention relates to a method of forming a reversible peptide-  
XX receptor complex that involves providing an immobilized receptor, and  
XX contacting the receptor with a zslg33 peptide (comprising residues 24-37  
XX of AAB62649), where the receptor binds to the zslg33 peptide. The method  
XX is useful for purifying cells, purifying a peptide, stimulating signal  
XX transduction in a cell expressing a receptor. It is also useful for  
XX modulating secretion of hormones, neural development and/or utilization,  
XX gastric contractility, nutrient uptake, secretion of digestive and  
XX pancreatic enzymes and hormones, secretion of insulin-like growth factor  
XX -I, secretion of non-zslg33 proteins. It is useful for modulating growth

CC hormone secretion in a mammal having a disease associated with abnormal  
CC levels of growth hormone, such as osteoporosis, bone repair, bone  
CC remodeling, low osteoblast levels, cartilage repair and remodeling,  
CC skeletal dysplasia, immune suppression, obesity, growth retardation,  
CC protein catabolic responses after surgery, cachexia, protein loss,  
CC dwarfism, wound healing and ovulation induction, treating a mammal having  
CC a metabolic disorder requiring neurological feedback, such as satiety  
CC regulation, glucose absorption and metabolism and neuropathy-associated  
CC gastrointestinal disorders, and stimulating glucose-induced insulin  
CC release in a mammal. The present sequence represents the short form of  
CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result  
CC from alternative splicing). GPR38 has homology to the human G-protein  
CC coupled receptor, GHS-R.

XX Sequence 386 AA;

Query Match 73.1%; Score 1575; DB 22; Length 386;  
Best Local Similarity 99.7%; Pred. No. 3.2e-140;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60  
Db 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60

QY 61 MLIGRYDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLCRSLVYGE 120

Db 61 MLIGRYDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLCRSLVYGE 120

QY 121 CTYATLHMTALSVERYLAICRPLRARVLTTRRRVRLAVLMAVALLSAGPFLVGE 180

Db 121 CTYATLHMTALSVERYLAICRPLRARVLTTRRRVRLAVLMAVALLSAGPFLVGE 180

QY 181 QDPGISVVGNGTARIASSPLASPPPLMSRAPPPSPGPEETAALFSECRPSPA 240

Db 181 QDPGISVVGNGTARIASSPLASPPPLMSRAPPPSPGPEETAALFSECRPSPA 240

QY 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGHQTVRL 300

Db 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGHQTVRL 300

RESULT 9

AAB68476  
ID AAB68476 standard; Protein; 271 AA.

XX AAB68476;

DT 23-JUL-2001 (first entry)

DE Amino acid sequence of a dog motilin receptor exon 1.

XX Motilin receptor; gastrointestinal disease; gastric motility disorder;  
XX gastroparesis; irritable bowel syndrome; diarrhoea.

XX Canis sp.

XX WO200132710-A1.

XX 10-MAY-2001.

XX 25-OCT-2000; 2000WO-US29426.

XX 29-OCT-1999; 99US-0162264.

XX (MERI) MERCK & CO INC.

XX Tan C, McKee K;

XX WPI; 2001-343479/36.

XX N-PSDB; AAF85447.

XX Novel polypeptides related to dog and rabbit motilin receptor  
XX polypeptide, comprising unique regions from dog and motilin receptor

PT amino acid sequence, useful for identifying compounds for treating  
 PT diarrhoea in humans -  
 PS Claim 4; Page 17; 42pp; English.  
 CC The present sequence is encoded by exon 1 of a dog motilin receptor gene.  
 CC The specification describes an unique sequence present in exon 1 of  
 CC the motilin receptor, which is not present in human or Spherooids  
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid  
 CC sequence is useful for measuring the ability of a compound to affect  
 CC motilin receptor activity. Motilin receptor polynucleotides and  
 CC polypeptides are used to identify therapeutic compounds which are  
 CC useful for treating gastrointestinal diseases and disorders such as  
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,  
 CC and diarrhoea.  
 CC  
 XX Sequence 271 AA;  
 SQ  
 Query Match 53.0%; Score 1143.5; DB 22; Length 271;  
 Best Local Similarity 76.7%; Pred. No. 1.1e-99;  
 Matches 230; Conservative 13; Mismatches 28; Indels 29; Gaps 4;  
 QY 1 MGSPNNGSDGEGARPPPALPCDERCSPPLGALVPTAVCLCFVVGSGNVMY 60  
 Db 1 MGGPNSSDGAEGAQ-----LPDERLCSPPLGALVPTAVCGLFVVGSGNLMY 53  
 QY 61 MLIGRYDMRTTNNLYGSMAYSDLLILGLPDLRYMRSPWVGPPLCRSLTYGEG 120  
 Db 54 LLIGRYDMRTTNNLYGSMAYSDLLILGLPDLRYMRSPWVGPPLCRSLTYGEG 113  
 QY 121 CTYATLHMTALSVRYALICPLPARVLTTRRRVALAVLMAVALSAGPPLFVGYE 180  
 Db 114 CTYATLHMTALSVRYALICPLPARVLTTRRRVALAVLMAVALSAGPPLFVGYE 173  
 QY 181 ODPGISVVGNGTARIRASPPPLWLSRAPPSPPSGPETAALAFSSECPSPA 240  
 Db 174 QDAG---GPGNGSRRLRA-----PSPPPGE---AALFSRECPSPS 211  
 QY 241 QLGALRWMLWTTAVFPLPCLSLILYGLIGELWSSRRRLGPAASGRGRHQTKRYL 300  
 Db 212 QLGALRWMLWTTAVFPLPCLSLILYGLIGELWSSRRRLGPAASGRGRHQTKRYL 271  
 RESULT 10  
 AAY54147  
 ID AAY54147 standard; Protein, 363 AA.  
 AC  
 XX AAY54147;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 XX The puffer fish motilin receptor protein encoded by clone 75E7.  
 DE  
 XX  
 XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;  
 KM Applied form; MTL-R1A; MTL-R1B; gastric motility disorder;  
 KM functional defect; neurological disorder; scleroderma; colonoscopy;  
 KM paraneoplastic syndrome; radiation induced dysmotility; diabetes;  
 KM infection; stress-related motility disorder; psychogenic disorder;  
 KM gastroparesis; gastro-oesophageal reflux disease; constipation;  
 KM chronic idiopathic pseudo obstruction; acute faecal impaction;  
 KM postoperative ileus; gallstones; infantile colic; diarrhoea;  
 KM irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;  
 KM endoscopy; duodenal intubation.  
 KM  
 XX Spheroides nephelus.  
 OS  
 XX  
 XX WO964436-A1.  
 PN  
 XX  
 PD 16-DEC-1999.  
 XX  
 XX 08-JUN-1999; 99WO-US12773.  
 PF  
 XX  
 PR 12-JUN-1998; 98US-0089098.

XX  
 PA (MERI) MERCK & CO INC.  
 XX  
 PI Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD,  
 XX Pong S, Smith RG;  
 DR N-PSDB; AAZ45405.  
 DR WPI; 2000-105868/09.  
 XX  
 PT Novel receptor protein for screening compounds used in treating  
 PT irritable bowel syndrome, constipation and other gastric conditions  
 PS  
 PS Claim 7; Fig 8; 44pp; English.  
 CC  
 CC The present sequence represents a motilin receptor (clone 75E7) which  
 CC is related to the human motilin receptor of the invention. The  
 CC specification describes a G-protein coupled receptor, designated  
 CC MTL-R1 (also GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see  
 CC AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a functional seven  
 CC transmembrane domain form, and MTL-R1B is a truncated five  
 CC transmembrane domain. The MTL-R1 proteins are used to identify agonists  
 CC and antagonists which can be used for treating gastric motility  
 CC disorders, functional defects, disorders secondary to neurological  
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced  
 CC dysmotility, diabetes, infections, stress-related motility disorders,  
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,  
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal  
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable  
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and  
 CC diarrhoea. They can also be used in the preparation for colonoscopy,  
 CC endoscopy and duodenal intubation.  
 CC  
 XX Sequence 363 AA;  
 SQ  
 Query Match 41.8%; Score 901; DB 21; Length 363;  
 Best Local Similarity 50.3%; Pred. No. 1.3e-76;  
 Matches 179; Conservative 56; Mismatches 77; Indels 44; Gaps 5;  
 QY 31 SPPPLGALVVTAVCLCFVVGSGNVMYVMTLIGRYDMRTTNNLYGSMAYSDLLILG 90  
 Db 31 SLFPASTLIPVTYICLIFVVGSGNVMYVMTLILIOYFKDMKTNNLYSSMAVSDLVIFLC 90  
 QY 91 LPEDLYRLMKYVPMFLGEGANCRILYHIFECCTSAITLHITLALSERLIALSPLRSKMV 150  
 Db 91 LPEDLYRLMKYVPMFLGEGANCRILYHIFECCTSAITLHITLALSERLIALSPLRSKMV 150  
 QY 151 TTRRRVALAVLMAVALSAGPPLFVGYEQDPGISVVGNGTARIRASPPPLW 210  
 Db 151 TTRRRVALAVLMAVALSAGPPLFVGYEQDPGISVVGNGTARIRASPPPLW 210  
 QY 211 SRAPPSPPSGPETAALAFSSECPSPAQL--GALRWMLWTTAVFPLPCLSLTYG 268  
 Db 186 ----HPDYNTG-----QCKHTGYAISGQHLHIMWSTYFCMCLLPLYG 229  
 QY 269 LIGRELWSSRRRLGPAASGRGRHQTKRYLVVLAFTICMLPFVHGRITTYINTDSR 328  
 Db 230 SIGCKLMKSKNDQGCALABERSHQTVLIVVLAFTICMLPFYIGNLTFQVQVDYD 289  
 QY 329 MMFSGYFNIVAQLYLASIPILYNLSKRYRAAFLLARKSRPGRFHSR 384  
 Db 290 TMLSLQNFWMASVWLCYLSASINPVVYNLSRKYRAAKRLFLHQ-RKPAHRGQ 344  
 RESULT 11  
 AAB68479  
 ID AAB68479 standard; Protein, 363 AA.  
 AC  
 XX AAB68479;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 XX Amino acid sequence of a motilin receptor polypeptide.  
 DE  
 XX



```

Db 122 VESCTYAVITLTTALSVETFAICPELRKAVVTKGRVAVLIVIAWVAFCSAGPIFVL 181
Qy 177 VGVGDDPGISVVPGLNGTARIASSPLASPPMLSRAPPPSPGPTAEAAALFSRECR 236
Db 182 VGVGHEH-----NGT-----DP--W-----DINECR 199
Qy 237 PS--PAQGLAKVLMVWTTAVAFPLCLSLYLGLIGRELMSSRRPLRGPAASGRERGHR 294
Db 200 PTEFAVRSGLLTVMVWSSIFPFLPVFCLTVLXSLIGRKLIMRRRRGDVAVGASLRDQNHK 259
Qy 295 QTKRVLVLLVLAFLICMLPPIHVRRIIYNEDS---RMVFSGYFNVALQPLFLSASIN 351
Db 260 QTKRVLVLLVLAFLICMLPPIHVRRIIYNEDS---RMVFSGYFNVALQPLFLSASIN 319
Qy 352 PLYVNLISKRYAAAFKLLARKSRPRGFRHSRDTAGEVAGDTGDTVGYTETSAN 407
Db 320 PLYVNLISKRYAAAFKLLARKSRPRGFRHSRDTAGEVAGDTGDTVGYTETSAN 365

```

## RESULT 13

AAy6293  
ID AAY6293 standard; Protein; 349 AA.

AC AAY6293;

DT 19-JUN-2000 (first entry)

DE A canine growth hormone secretagogue receptor (GHSR) protein.

KM Canine; growth hormone secretagogue receptor; GHSR; dog.

OS Canis sp.

PN WO200009538-A2.

PD 24-FEB-2000.

PF 06-AUG-1999; 99WO-US17915.

PR 10-AUG-1998; 98US-0095960.

PA (MERI) MERCK & CO INC.

PI Howard AD, Palyha OC, Smith RG, Tan CP;

DR WPI; 2000-224272/19.

DR N-PSDB; AAZ61492.

PT A novel canine growth hormone secretagogue receptor used to identify agonists and antagonists

PS Claim 6; Fig 2; 32pp; English.

CC The present sequence represents a canine growth hormone secretagogue receptor (GHSR). The GHSR cDNA sequence was isolated from a canine cDNA library constructed from spleen. The GHSR polynucleotide is a source of probes and primers, which can be used to isolate GHSR genes from other species. It is also used for recombinant expression of the polypeptide. The GHSR polypeptide is used in assays to identify agonists, antagonist and inhibitors. Such identified compounds can serve as leads for the development of veterinary pharmaceuticals that can be used to treat dogs having illnesses in which inappropriate activity of the canine GHSR proteins involved. The polypeptides may also be used to raise antibodies for diagnostic and therapeutic purposes.

CC Sequence 349 AA;

Query Match 39.94; Score 859.5; DB 21; Length 349;  
Best Local Similarity 46.24; Pred. No. 1e-72;  
Matches 166; Conservative 50; Mismatches 98; Indels 69; Gaps 8;

```

Qy 14 AREPE-----WPALPCDEBRRCSPFLGALVPTAVACLCFVVGSGNVVTMLIGRDM 69
Db 6 AREPGSAGMDLP-----LFAPILAGVTATCAVCLFAVGAGNLLTLVRRFREEL 57
Qy 70 RTTNLYIGSNVAVSDLLILGLPFDLYRLMRSRPWFGLPCLSLYVGEQTYATLLHM 129
Db 58 RTTNLYIGSNVAVSDLLILGLPFDLYRLMRSRPWFGLPCLSLYVGEQTYATLLHM 117
Qy 130 TALVERVLAICPLRAVAVVTRRRVALLIYALVALISAGPFLPLVGVGDDPGISVVP 189
Db 118 TALVERVLAICPLRAVAVVTRRRVALLIYALVALISAGPFLPLVGVGDDPGISVVP 170
Qy 190 GLNGTARIASSPLASPPMLSRAPPPSPGPTAEAAALFSRECRPS--PAQGLAV 247
Db 171 -NGT-----DPRD-----TRCRATEFVVRSGLLTA 195
Qy 248 MLWVTTAVFPLPCLSLYLGLIGRELMSSRRPLRGPAASGRERGRQTKRVLVLLVLA 307
Db 196 MWMVSVFFFLPVFCLTVLYGLIGRKLIMRRRRGDVAVGASLRDQNHK 255
Qy 308 IICMLPPIHVRRIIYNEDS---RMVFSGYFNVALQPLFLSASINPLYNLISKRYA 364
Db 256 IICMLPPIHVRRIIYNEDS---RMVFSGYFNVALQPLFLSASINPLYNLISKRYA 315
Qy 365 AAFLLTARKSRPRGFRHSRDTAGEVAGDTGDTVGYTETSAN 407
Db 316 AVFLLIGFEPFSQKRLTLWDESSR-----AMTESSIN 348

```

## RESULT 14

AAy90632  
ID AAY90632 standard; Protein; 366 AA.

AC AAY90632;

DT 21-AUG-2000 (first entry)

DE Human G protein-coupled receptor GHSR.

KM G protein-coupled receptor; GPCR; constitutively active;

KW intracellular loop 3; transmembrane domain 6; drug screening;

XX agonist; antagonist.

OS Homo sapiens.

PN WO200022129-A1.

PD 20-APR-2000.

PF 12-OCT-1999; 99WO-US23938.

PR 13-OCT-1998; 98US-0170496.

PA (AREN-) ARENA PHARM INC.

PI Behan DP, Chalmers DT, Liaw CW;

DR WPI; 2000-329165/28.

DR N-PSDB; AAA30643.

PT Non-endogenous constitutively activated human G protein-coupled

PT receptors, useful for identifying agonists for use as pharmaceutical

PS Example 1; Page 167-169; 341pp; English.

CC The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-CC and AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-AA30743-CC and AAA30775-AA30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (Tm6). A non-endogenous amino acid, X, is substituted for an endogenous residue in IC3 at a position 16 amino

CC acids N-terminal of an endogenous proline in TM6 to form a sequence  
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg  
 CC or Ala, and is preferably Lys. When the endogenous residue at this  
 CC position is Lys, this residue is replaced by His, Arg or preferably Ala.  
 CC The 15 amino acid stretch between the substituted amino acid and the Pro  
 CC may be endogenous, non-endogenous, or a mixture of endogenous and  
 CC non-endogenous residues. The constitutively active GPCRs are useful for  
 CC identifying antagonists, agonists and partial agonists for use as  
 CC pharmaceutical agents. The mutant proteins are also useful in research  
 CC settings for elucidating the roles of the receptors in normal and  
 CC diseased conditions. Antagonists for a particular GPCR are useful for  
 CC treating diseases and disorders associated with that receptor. Because  
 CC the novel mutant GPCRs are constitutively active, they can be used  
 CC directly for screening of compounds without the need for endogenous  
 CC ligands. The present sequence represents a human wild-type GPCR referred  
 CC to in an exemplification of the invention.

XX Sequence 366 AA;

Query Match 39.8%; Score 857.5; DB 21; Length 366;

Best Local Similarity 44.5%; Pred. No. 1.7e-72;

Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

DB 5 WNGSDPEGA-----REPPPALPPCD---ERRCSPPGALVPTAVCLCTFVGVSGN 56  
 2 WNAITBEEEGFNLTADLDMDASPGNDSIGDELQFPAPLAGVATCVALFVVGIAQN 61

QY 57 VVTVMILGRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLMRSPWVFGPLCLSLY 116  
 62 LITMLVVSFRRLRTTNLYLSSMAFSDDLIFLCMPDLVRLMQYRPMNFGDLCLCFQF 121

QY 117 VEGGCTATLLHMTALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 176  
 122 VSESCYATVLTITLALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 181

QY 177 VGEODPGISVVGNGLTARIASPPPLMSRAPPPSPSGPETAAMALFSRECR 236  
 182 VGVGHE-----NGT-----DP--W-----DTNECR 199

QY 237 PS--PAQLGALRYMLWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294  
 200 PTEFAVRSGLLTVMWVSSIFFLPVCLTVLSTLGRKLMRRRRGDVAVGASLRDQNHK 259

QY 295 QTKRVLIVVLAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 351  
 260 QTKMVAIVVFAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 319

QY 352 PLYNLSKYYRAAFAKLLARKSRPGRHRSRDTAGEVAGDTGDTVGYTETSAN 407  
 320 PLYNLSKYYRAVAFRLIGFEPFSQKSLTKDESSR-----AWTESSIN 365

DB

RESULT 15  
 AAY70345  
 ID AAY70345 standard; Protein; 366 AA.

AC AAY70345;

DT 20-JUN-2000 (first entry)

DE Human G protein-coupled orphan receptor, GHSR.

XX G protein-coupled orphan receptor; GPCR; agonist; G protein; treatment;

KW GPCR fusion protein; inverse agonist; drug; GHSR; human.

OS Homo sapiens.

PN W0200006597-A2.

PD 10-FEB-2000.

PF 30-JUL-1999; 99WO-US17425.

XX

PR 31-JUL-1998; 98US-0094879.  
 PR 30-OCT-1998; 98US-0106300.  
 PR 04-DEC-1998; 98US-0110906.  
 PR 26-FEB-1999; 99US-0121651.

PA (AREN-) ARENA PHARM INC.

PI Behan DP, Chalmers DT, Liaw C, Lin I, Lowitz K, Chen R;

DR WPL; 2000-195260/17.

DR N-PSDB; AAZ51463.

PT Identification of a compound useful as a therapeutic agent, comprises  
 PT identifying a compound against constitutively activated G  
 PT protein-coupled orphan receptors

PS Example 2, Page 123; 123pp; English.

XX The patent discloses a method of identifying agonists and inverse or  
 CC partial agonists to the endogenous, constitutively activated  
 CC G protein-coupled orphan receptors (GPCRs), by contacting them with a  
 CC GPCR fusion protein comprising a GPCR and a G protein. Determining  
 CC expression of GPCRs in tissue samples can be used to identify related  
 CC diseases. Inverse agonists to these receptors can be used as drugs for  
 CC treating GPCR-related diseases. The present protein sequence is that of  
 CC human G protein-coupled orphan receptor, GHSR.

Sequence 366 AA;

Query Match 39.8%; Score 857.5; DB 21; Length 366;

Best Local Similarity 44.5%; Pred. No. 1.7e-72;

Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

DB 5 WNGSDPEGA-----REPPPALPPCD---ERRCSPPGALVPTAVCLCTFVGVSGN 56  
 2 WNAITBEEEGFNLTADLDMDASPGNDSIGDELQFPAPLAGVATCVALFVVGIAQN 61

QY 57 VVTVMILGRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLMRSPWVFGPLCLSLY 116  
 62 LITMLVVSFRRLRTTNLYLSSMAFSDDLIFLCMPDLVRLMQYRPMNFGDLCLCFQF 121

QY 117 VEGGCTATLLHMTALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 176  
 122 VSESCYATVLTITLALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 181

QY 177 VGEODPGISVVGNGLTARIASPPPLMSRAPPPSPSGPETAAMALFSRECR 236  
 182 VGVGHE-----NGT-----DP--W-----DTNECR 199

QY 237 PS--PAQLGALRYMLWTTAYFFLPFLCLSLYGLIGRELWSSRRPLRGPAASGRGRHR 294  
 200 PTEFAVRSGLLTVMWVSSIFFLPVCLTVLSTLGRKLMRRRRGDVAVGASLRDQNHK 259

QY 295 QTKRVLIVVLAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 351  
 260 QTKMVAIVVFAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 319

QY 352 PLYNLSKYYRAAFAKLLARKSRPGRHRSRDTAGEVAGDTGDTVGYTETSAN 407  
 320 PLYNLSKYYRAVAFRLIGFEPFSQKSLTKDESSR-----AWTESSIN 365

Search completed: March 16, 2003, 15:03:58

Job time: 42 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2003, 00:44:59 ; Search time 3287 Seconds

(without alignments)  
10051.352 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040

Sequence: 1 atgggagagccctggagacg.....acgtgagagcagatgggataa 2040

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrdl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_esthum:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.2	28.7	590	12	BF513101 UT-H-BM1-
2	255.8	12.5	425	17	AQ138681 HS_3073_A
3	226.8	11.1	500	12	BF603623 BF603623
4	142	7.0	504	13	AZ451922 AZ451922
5	127	6.2	855	17	BI757121 BI757121
6	127	6.2	1010	14	BM925480 BM925480

7	125.4	6.1	634	12	BG714306
8	125.4 <td>6.1</td> <td>1055</td> <td>14</td> <td>BM808829</td>	6.1	1055	14	BM808829
9	125.2	6.1	540	14	BM737725
10	124.8	6.1	1027	13	BM543497
11	124	6.1	1144	13	BM546496
12	122.8	6.0	941	17	CNS036WH
13	119	5.8	1603	14	BM808823
14	117.8	5.8	738	13	BI600170
15	111.8	5.5	928	13	BI758259
16	111.2	5.5	933	14	BQ72315
17	109.6	5.4	448	12	BF323227
18	109.2	5.1	828	13	BI597845
19	104.8	5.0	980	13	BM543468
20	102.6	5.0	1189	11	AK005368
21	102.2	4.8	785	13	BI754749
22	97	4.8	843	17	CNS04SGU
23	97	4.8	670	17	AG046172
24	96.8	4.7	720	13	BI753905
25	96.8	4.7	555	12	BE751626
26	96.6	4.7	825	9	AU079556
27	96.6	4.7	1115	13	BM545572
28	94	4.6	959	17	CNS04GXP
29	93.6	4.6	961	14	BQ056144
30	92.2	4.5	1139	14	BM922349
31	92.2	4.5	1798	17	AG171124
32	91	4.5	460	10	AW206493
33	90.8	4.4	624	10	BS587450
34	89.6	4.4	326	10	BE654108
35	89.2	4.4	462	10	AW138148
36	89.2	4.4	544	9	AI344017
37	89.2	4.4	741	13	BI224313
38	89	4.4	984	9	AL550664
39	88.8	4.4	961	17	CNS030RC
40	88.4	4.3	975	17	CNS039RV
41	88.4	4.3	866	9	AL553958
42	88	4.3	1083	14	BM805382
43	87.8	4.3	1072	14	BQ055067
44	87.6	4.3			
45					

## ALIGNMENTS

RESULT 1  
LOCUS BF513101 590 bp mRNA linear EST 07-DEC-2000  
DEFINITION UT-H-BM1-amm-e-07-0-UT-sl CGAP\_Sub7 Homo sapiens cDNA clone  
IMAGE:3070549 3', mRNA sequence.

ACCESSION BF513101  
VERSION BF513101.1 GI:11598280  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 590)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [CGAPbs-r@mail.nih.gov](mailto:CGAPbs-r@mail.nih.gov)  
Oligo-dT track not found. Not 1 site shown in beginning of sequence  
is likely internal to the message. cDNA library preparation: M.B.  
Soares Lab Clone distribution: NCI-CGAP clone distribution  
Information can be found through the I.M.A.G.E. Consortium/LNL at:  
[www.bio.lnh.gov/bbrp/image/image.html](http://www.bio.lnh.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=No.

FEATURES  
source 1.590  
/organism="Homo sapiens"

```

/db xref="taxon:9606"
/clone="IMAGE:3070549"
/clone_1ib="NCI CGAP Sub7"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7
is a subtracted library derived from NCI CGAP Sub6. The
NCI CGAP Sub7 library had 12 million recombinants. A
single-stranded DNA preparation of NCI CGAP Sub6 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI CGAP Kid3 pool 1 L1AM
1334-3337, 3682-3683, 3798-3803 (IMAGE Clones
132376-132391, 145608-145675, 150052-150285);
NCI CGAP Kid5 pool 1 L1AM 3338-3342, 3722-3725, 3776-3778
(IMAGE Clones 132392-132531, 1471368-1472903,
1492104-1493255); NCI CGAP L45 pool 1 L1AM 3375-3582,
3851-3854 (IMAGE Clones 1414920-1417991, 1520904-1522439
); NCI CGAP G4 pool 1 L1AM 3164-3167, 3716-3720,
3733-3735 (IMAGE Clones 1257096-1258631, 1469064-1470983
1475593-1476743); NCI CGAP Pr2 pool 1 L1AM 2457-2459,
2758-2759, 3062-3068 (IMAGE Clones 985608-986759,
1101192-1101959, 1217928-1220615); NCI CGAP C10 pool 1
L1AM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255
, 114584-1145351). (6% of the driver population), plus a
pool of 3,840 arrayed clones from NCI CGAP Sub1 (IMAGE
Clones 2708616-2710535) and NCI CGAP Sub2 (IMAGE
Clones 2710536-2712455) (4% of the driver population
), plus a pool of 11,136 clones from NCI CGAP Sub3 (IMAGE
Clones 2712456-2723591) (10% of the driver population),
plus a pool of 5,472 clones from NCI CGAP Sub4 (IMAGE
Clones 2723592-2729326) (40% of the driver population),
plus a pool of 4032 clones from NCI CGAP Sub6 (IMAGE
Clones 2728969-2733190) (40% of the driver population).
Subtraction was performed as previously described [Bonaldo
, Lennon & Soares (1996): Normalization and Subtraction:
Two Approaches To Facilitate Gene Discovery. Genome
Research 6, 791-806.
TAG SEQ=None found"

```

BASE COUNT 130 a 173 c 157 g 130 t

ORIGIN

Query Match 28.7%; Score 585.2; DB 12; Length 590;  
Best Local Similarity 99.5%; Pred. No. 6.6e-119;  
Matches 587; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

835 CGGCGCGCTGCGAGGCGCGCGCGCTCGGCGCGGAGAGAGCGCCGCGCAACCAAGCGC 894  
1 CGGCGCGCTGCGAGGCGCGCGCGCTCGGCGCGGAGAGAGCGCCGCGCAACCGTCCGC 60  
895 GTCTCGGTAAGTGGAGCGCGCGTGTTCAGAAAGCGCTCTGCACTGCGCCCGCGC 954  
61 GTCTCGGTAAGTGGAGCGCGCGTGTTCAGAAAGCGCTCTGCACTGCGCCCGCGC 120  
955 GGAACCGCGCAAGCGTGGTCCCTTCCTGCTGCGCCAGCTCTGGGCGCGCTTCCA 1014  
121 GGAACCGCGCAAGCGTGGTCCCTTCCTGCTGCGCCAGCTCTGGGCGCGCTTCCA 180  
955 GGAACCGCGCAAGCGTGGTCCCTTCCTGCTGCGCCAGCTCTGGGCGCGCTTCCA 1014  
1015 GCTGCCCTTCTTATTTGATTCAGACCTCCACCGCGGTACTTCCATCCCGCAGAAA 1074  
181 GCTGCCCTTCTTATTTGATTCAGACCTCCACCGCGGTACTTCCATCCCGCAGAAA 240  
1075 ACCATGTCCTGTCGCCAGAGCTCTGGGGAACCCAGGCGCTTTGAGGTTGATCCC 1134  
241 ACCATGTCCTGTCGCCAGAGCTCTGGGGAACCCAGGCGCTTTGAGGTTGATCCC 300  
955 GGAACCGCGCAAGCGTGGTCCCTTCCTGCTGCGCCAGCTCTGGGCGCGCTTCCA 1014  
1135 CGATTCGATTCAGTACAGAGCTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGT 1194  
301 CGATTCGATTCAGTACAGAGCTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGT 360  
1195 TGGTAATTTTAAATCAACCAAGCTTTAGATGCCAATAGAGAGTCTCAGAGTCTC 1254  
361 TGGTAATTTTAAATCAACCAAGCTTTAGATGCCAATAGAGAGTCTCAGAGTCTC 420

```

QY 1255 TTGAGAGAGAGAGAGAGATTTCAATTAAGTAAATTTTATTTAATGTAAGTATGCT 1314
DB 421 TTGAGAGAGAGAGAGAGATTTCAATTAAGTAAATTTTATTTAATGTAAGTATGCT 480
QY 1315 GAAGGCTTAAGTAACTTGTCTATCAAAAGTAAAGTATGACAGCTGTGTGA 1374
DB 481 GAAGGCTTAAGTAACTTGTCTATCAAAAGTAAAGTATGACAGCTGTGTGA 540
QY 1375 ATTCTTTCAACAGAGAAAGAAAGTCTCCGAGAGTGGTGTGTGA 1424
DB 541 ATTCTTTCAACAGAGAAAGAAAGTCTCCGAGAGTGGTGTGTGA 590

```

RESULT 2 425 bp DNA linear GSS 24-SEP-1998  
A0138681  
LOCUS  
DEFINITION  
HS\_3073\_A2\_D02\_MR\_CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate=3073 Col=4 Row=G, DNA sequence.  
ACCESSION  
A0138681  
VERSION  
A0138681.1 GI:3529334  
KEYWORDS  
GSS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 425)  
Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holman T.,  
Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and  
Hood L.  
Sequence-tagged connectors: A sequence approach to mapping and  
screening the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
CONTACT: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3073 row: G column: 4  
Class: BAC ends  
High quality sequence stop: 425.  
Location/Qualifiers  
1..425  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate=3073 Col=4 Row=G"  
/clone\_1ib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
E-Coli DH10B"

BASE COUNT 86 a 91 c 111 g 133 t 4 others

ORIGIN

Query Match 12.5%; Score 255.8; DB 17; Length 425;  
Best Local Similarity 94.2%; Pred. No. 4.1e-46;  
Matches 274; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 1371 TGAATTTCTTTCAACAGAGAAAGTCTGTCGGAAGGAGGTTGTGAGAGAG 1430  
DB 51 TGAATTTCTTTCAACAGAGAAAGTCTGTCGGAAGGAGGTTGTGAGAGAG 110  
QY 1431 CCTGCAAGAGCGGCTTTGTCAGAGAAATGCTCTGATTATGTCAGCTTGATAA 1490  
DB 111 CCTGCAAGAGCGGCTTTGTCAGAGAAATGCTCTGATTATGTCAGCTTGATAA 170  
QY 1491 CACATATGGAGAGCTTATGAGAGTTTAAAGCAAGTATCCAGCTGAGCTTGCT 1550  
DB 171 CACATATGGAGAGCTTATGAGAGTTTAAAGCAAGTATCCAGCTGAGCTTGCT 230  
QY 1551 CATTTTCTGGGAGTACAGATCTGCTAGTGAAGTTTCTCAATTAATTTTGTGCTT 1610

Db 231 CATTTCCTGGGGGAGATGCTGCTTANAGAGTTTCTTAATTAATTGCTG-T 289

Qy 1611 ACTTGTATTCAGATGATGCTTGTTCGGGGGATTAATTTGCTTC 1661

Db 290 ACTTGTATTCAGATGATGCTTGTTCGGGGGATTAATTTGCTTC 340

RESULT 3  
BF603623 500 bp mRNA linear EST 25-APR-2001  
LOCUS 269181 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION BF603623  
VERSION BF603623.1 GI:11701421  
KEYWORDS EST.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 500)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Caes,B., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,  
Perteaux,G., Holt,I., Karaycheva,S., Liang,F., Quackenbush,J. and  
Keeler,J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
2118013  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -mnscore 18  
and -mnmult 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCAGGACGACGACG  
Plate: 49 row: 1 column: 6  
Seq primer: ATTTAGGTGACACTATAG.  
Location/Qualifiers  
1..500  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;  
library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinosus muscle, and fetal  
longissimus muscle."

BASE COUNT 71 a 179 c 139 g 111 t

ORIGIN

Query Match 11.1%; Score 226.8; DB 12; Length 500;  
Best Local Similarity 68.2%; Pred. No. 1.1e-39;  
Matches 315; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 92 CGCCCTTCCCTGGGGGCGCTGCGCCGTCACCCGCTGTGCTGCTGCTGCTG 151

Db 39 CCGCTTCCCGCGCGCTGCTGCGGGGTCACACCCACCTCGGCGCTCTTGATG 98

Qy 152 TCGGGGTAGCGGCAACGTCGTGATCGTATCGGCGCTACCGGACATCGGGA 211

Db 99 TGGGATGCGGGCACTGCTCACCATCTGTGTGTGCGCTTCCGGAGACTCGTA 158

Qy 212 CCACCACTTGTGCTGCGGACGATGCGGTGTCGACCTACTCATCTGCTCGGG 271

Db 159 CCACCACTTGTGCTGCGGACGATGCGGTGTCGACCTACTCATCTGCTCGGA 218

Qy 272 TGCCGTTGACCTGTAACCGCTCTGCGCGCTGAGGCTTGGGCGCTGCTCT 331

Db 219 TGCCCTGACCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCT 278

Qy 332 GCGGCTGCTCCCTTACGTCGCGGCGGCTGACCTACGCAACGCTGTCATGAC 391

Db 279 GCAAACTCTTCAGTTTTCAGCGAGAGCTGACCTACGCTACGCTGCTCACC 338

Qy 392 GCGTCAGCTGAGCGCTACCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 451

Db 339 GCGTCAGCTGAGCGCTACCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 398

Qy 452 CCGGCGCGCGCTGCGCGCTGCGCTGCGCTGCGCGCTGCGCGCTGCGCGCTG 511

Db 399 CCAAGGCGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 458

Qy 512 GTCCCTTCTTCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 553

Db 459 GCGCCATCTTCTGCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 500

RESULT 4  
AZ451922/c 504 bp DNA linear GSS 04-OCT-2000  
LOCUS 1M0251C06R Mouse 10kb plasmid UNGC1M library Mus musculus genomic  
ACCESSION AZ451922  
VERSION AZ451922.1 GI:10608187  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muride; Murinae; Mus.  
1 (bases 1 to 504)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.  
and Wright,D., Weis,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weis  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0251 row: C column: 06  
Seq primer: CACACGGAACAGCTATGAC  
Class: plasmid ends  
High quality sequence file: 504.  
Location/Qualifiers  
1..504  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="UNG1M0251C06"  
/clone\_lib="Mouse 10kb plasmid UNGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42ny. Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The

adapored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 156 a 110 c 107 g 131 t  
ORIGIN

Query Match 7.0%; Score 142; DB 17; Length 504;  
Best Local Similarity 75.6%; Pred. No. 6e-21;  
Matches 189; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 1682 GTGCTGTGCTTATGTTGACAGTGTGTGTTCTGGCATTATTAATTTGC-TGTTGCC 1740  
DB 270 GTCTTGTTCGAGCGATGTGTGAGTTCTGGCATTCAATTTGCTTGTGCTT 211  
QY 1741 TTCCAGTTGGCGATCATTTACATATAACGGAATTCGGGATGATGATCTTCT 1800  
DB 210 TTCCAGTTGGCGAATCTTTTCATGACACATAAAATTCGGGATGATGATCTTCT 151  
QY 1801 CAGTACTTAACTCGTGTGCTGCACTTTCTATCTGAGCGCATCTATCAACCAATC 1860  
DB 150 CAGTACTTAACTGATGATCTTCTATACCGAGTGCATCATCAAGCCATC 91  
QY 1861 CTCTACACCTCATTTCAAGAGTACAGAGCGGCTTTAACTGCTGCTGCAAG 1920  
DB 90 CTCTACACCTCATGTCATTAATGACAGAGCTGTCAGAAATGCTTCTGCGAGA 31  
QY 1921 AAGTCCAGGC 1930  
DB 30 CAGTCCAGGC 21

## RESULT 5

B1757121 855 bp mRNA linear EST 25-SEP-2001  
LOCUS B1757121  
DEFINITION 603030882F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5200868 5',  
mRNA sequence.

ACCESSION B1757121  
VERSION B1757121  
KEYWORDS B1757121.1 GI:15748699  
SOURCE EST.  
ORGANISM human.

REFERENCE Homo sapiens  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
JOURNAL

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
plate: LLM1503 row: d column: 21  
High quality sequence stop: 803.

## FEATURES

source  
1. 855  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5200868"  
/clone\_lib="NIH\_MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site\_1: Nct1;

Site 2: EcoRV (destroyed), RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is 0.190-0T primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

BASE COUNT 103 a 300 c 299 g 153 t  
ORIGIN

Query Match 5.2%; Score 127; DB 13; Length 855;  
Best Local Similarity 55.2%; Pred. No. 1.3e-17;  
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGGGCGGAGCCCGCTGGCCCGCGCTGCGCCCTTGGACAGCGCG 86  
DB 60 CAGCAGCCCCGCGCCCGCCCGCGCGCGCTCCACACCCGGGGCTGAGCTGAGCGCGCGCT 119  
QY 87 CTGCTGCTTTTCCCTGCGGGCGCTGTGCGGCTGACCGCTGTGTGCTGTCTT 146  
DB 120 GGGCGTGACACTGCGCTCTGTGGCCAGAGTGTCTTCAACCGGCTTACGCACTATCTG 179  
QY 147 GCTGTGCGGGTGAAGCGGCAACGTG-----GTGACCGGATGATGAGGGGCTAACG 200  
DB 180 GCGCTGGGCGGGGCGGCAATGCGCTGTCCGTGACGTGTGTGAAGCGCGCGCG 229  
QY 201 GACATGCGGACCAACCACTTGTACTGCGGCAAGATGCGCGTCCGACTACTCAT 260  
DB 240 GCGCGCGGGGCGCGCGCGCGCACACAGTGTCAAGCTGCGGCGCTGCTGTCTGCT 299  
QY 261 CTGTCTGCGGCTGCGGCTGTGACCTGTACCGCTTGTGGCGCTGCGG---CCCTGGGTGTT 317  
DB 300 GCTGTGCGGCGCGGCGGTGAGCTTCAACAGCTTGTGTCTCACTACCTCTGGGCTT 359  
QY 318 CGGGCGCGGCTGTGCGCGCTGTCCCTCAAGTGGGCGAGCGCTGACCTAGCGCAAGCT 377  
DB 360 CGGCACTTGGGCTGTGCGGCTGTACTTCTGTGACAGAGTGTGGCTAGCGCAAGCT 419  
QY 378 GGTGACATGACCGCGCTGAGCGCTGAGCGCTACCTGCGCATGTGCGCGCTCGCGC 437  
DB 420 GTTGAAGCTGTGAGCGCTGAGCGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCT 479  
QY 438 CCGGCTTTTGTGACACCGCGCGCGCTCGCGCGCTCATGCTGTGCTGTGGGCGGCTG 497  
DB 480 CCGGCGCGCTGTGAGCGCAACCGCGCGCGCTGTGCGCTGTGCGCTGTGCGCGCTG 539  
QY 498 GCTGCTCTGCGCGGCTCTTCTTGTCTGTGCGCGCTGTGAGCGAGACCGCG 550  
DB 540 GTTGGCTTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 592

## RESULT 6

BW925480 1010 bp mRNA linear EST 12-MAR-2002  
LOCUS BW925480  
DEFINITION AGENCOURT.6625013 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5763267  
5', mRNA sequence.

ACCESSION BW925480  
VERSION BW925480.1 GI:19375859  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

REFERENCE Homo sapiens  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
JOURNAL

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM12814 row: n column: 04  
 High quality sequence stop: 703.

## FEATURES

source

Location/Qualifiers  
 1..1010  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5763267"  
 /clone\_1ib="NIH\_MGC\_114"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dr  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Instituto de Genética, Research Genetics tracking code 019. Note:  
 this is a NIH\_MGC Library.)"

BASE COUNT 149 a 348 c 337 g 176 t  
 ORIGIN

Query Match 6.2%; Score 127; DB 14; Length 1010;  
 Best Local Similarity 55.2%; Pred. No. 1.4e-17;  
 Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

27 CGGCCCCGAGGGGGGCGGAGAGCGCCGCGCTGCGCCGCTTGGACGAGCGCG 86  
 Db CAGCAGCCCGGGGGCCCGCGGCGCCAGCTCCAAACCGGGGCTAGCGTACCGCGCT 141  
 QY 87 CTGCTCCGCTTTTCCCTTGGGGGCGCTGTGCGCGGTGACCGCTGTGCTGTG 146  
 Db 142 GGGGTGACACTCCCTCTGGGCAAGGTGCTGTACCGCGCTTACGCACTATCTG 201  
 QY 147 CGTGTGCGGGGAGCGGCAAGCTG-----GTGACCGGTATGCTATGGGGGCTAACG 200  
 Db 202 GGGCTGGGGCGGGGCGGCAATGCGCTGTCCGTGACAGTGTGTGAAGCGGGCGCG 261  
 QY 201 GGACATGCGGACCAACCAACTTGTACTGTGGGAGCATGCGGTGCGCACTTACTAT 260  
 Db 262 GCGCGGGGGGCGGCGGCGGCAACAGTGTGACCTGCGGCGCTGCGGCGCTGTGCT 321  
 QY 261 CTGCTGGGCTGCGCTTGTGACCTGTACCGCTTGGCGCTGCGG--CCCTGGGTGT 317  
 Db 322 GCTGTGGCGGTGCGGTGAGCTTACAGCTTGTGTGTCACTACCTCGGGTCTT 381  
 QY 318 CGGCGCGTGTGCTGCGCGCTGTCCCTCTAAGTGGGCGAGGGCTGCACTAGCCACGCT 377  
 Db 382 CGGCGACTTGGGCTGCGCGGCTTACTTGTGTGACAGCTGTGCGCTTACGCCACGCT 441  
 QY 378 GCTGACATGACCGCGCTCAGCGTGTGAGCGCTACTGTGCGCATCTGCGCGCGCT 437  
 Db 442 GCTGAGGTGAGCGGCTGAGCGGCGGAGCGGTGCTTACGCGGTGTGCGCACTGCTG 501  
 QY 438 CCGGCTTGTGTACACCGGCGCGCGCTGCGCGCTTACGTGTGCTTGGGCGGTGCG 497  
 Db 502 CCGAGCTGTGTGACCGCACCGCGGACCGGCTGTGTGTGCGCTGTGTGGCGCGCTC 561  
 QY 498 GCTGCTCTTGTGCGCGCTTCTGTGTCTGTGTGGGCGGTGAGGAGGAGACCGCG 550  
 Db 562 GCTGCGCTGCGCTGCGGCTGCGGCTGCTATCATGTGGGAGGAGGAGGAGCTGCG 614

RESULT 7  
 BG714306 634 bp mRNA linear EST 08-MAY-2001  
 LOCUS BG714306 602669812F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:4792730 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG714306  
 VERSION BG714306.1 GI:13993237

## KEYWORDS

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 634)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10671 row: g column: 03

High quality sequence stop: 633.

Location/Qualifiers

1..634

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4792730"

/clone\_1ib="NIH\_MGC\_96"

/tissue\_type="hypothalamus"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pBluescripter (modified

pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcagag

); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to R0.5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 74 a 221 c 243 g 96 t

ORIGIN

Query Match 6.1%; Score 125.4; DB 12; Length 634;

Best Local Similarity 55.0%; Pred. No. 2.9e-17;

Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

27 CGGCCCCGAGGGGGGCGGAGAGCGCCGCGCTGCGCCGCTTGGACGAGCGCG 86

Db 49 CAGCAGCCCGGGGCCCGCGGCGCCAGCTCCAAACCGGGGCTAGCGTACCGCGCT 108

QY 87 CTGCTCCGCTTTTCCCTTGGGGGCGCTGTGCGGTGACCGCTGTGCTGTGCT 146

Db 109 GGGCTGTGACACTCGCTTGGGCGCAAGGTGCTGTGTACCGCGCTTACGCACTATCTG 168

QY 147 CGTCTGCGGGGTGAGGGGCAAGCTG-----GTGACCGGTATGCTATCGGGGCTAACG 200

Db 169 GGGCTGGGGCGCGGCGGCAATGCGCTGTGCGCAAGTGTGTGAAGCGCGGGCGG 228

QY 201 GGACATGCGGACCAACCAACTTGTACTGTGGGAGAGATGGCGGTGCGACCTACTAT 260

Db 229 GCGAGGGGGGCGGCGGCGGCAACAGTGTGACCTGTGCGCGCTGCGGCGCTGCTGCT 288

QY 261 CTTGCTGGGCTGCGCTTGTGACCTGTACCGCTTGTGCGCGCTGCGG--CCCTGGGTGT 317

Db 289 GCTGTGCGGCTGCGGCTGTGAGCTTACAGCTTGTGTGTGTGTGTGTGTGTGTGT 348

QY 318 CGGCGCGCTGCTTGTGCGCGCTGTCCCTCTAAGTGGGCGAGGGCTGCACTTACGCACT 377

Db 349 CAGCGACTTGGGCTGCGGCGGCTTACTTGTGTGACAGCTGTGCGGCTTACGCCACGCT 408

QY 378 GCTGACATGACCGCGCTCAGCGTGTGAGCGCTACTGTGCGCATCTGCGCGCGCTGCGCG 437

Db 409 GCTGAGCGTGTGAGGCGCTTACGCGCGGAGGCGGCTTACGCGGTGTGTGTGTGTGTGT 468





CNMA Library Preparation: Life Technologies, Inc.  
 CNMA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM12726 row: e column: 02  
 High quality sequence start: 27  
 High quality sequence stop: 519.  
 Location/Qualifiers

## FEATURES

source

1..1144  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5729257"  
 /clone\_1ib="NH\_MGC\_124"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SORT6; Site 1: EcorV  
 (destroyed); Site 2: NotI; RNA source male hippocampus,  
 age 27. Library is oligo-dT primed and directionally  
 cloned (EcorV site is destroyed upon cloning). Average  
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 012."

BASE COUNT 188 a 384 c 371 g 200 t 1 others  
 ORIGIN

Query Match 6.1%; Score 124; DB 13; Length 1144;  
 Best Local Similarity 55.2%; Pred. No. 6.5e-17;

Matches 288; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

OY 27 CGGCCCCGAGGGGGCGGGAGCCCGCTGCGCCCGCTTGCAGCGAGCGCG 86  
 DB 67 CAGCAGCCCCCGGCGCTGCGGCCAGCTCCAGCCGGGCTGAGCGCTGAGCGCT 126  
 OY 87 CTGCTGCGCTTCCCTGCGGGCGCTGAGTGCAGTGCCTGAGTGCCTGCTGCT 146  
 DB 127 GGGCGTGGACATCGCTTCTGGGCGAAGTCTGTTCAACCGCGCTTACGCACTATCTG 186  
 OY 147 CGTGTGCGGGGTGAGCGGCAACGTG-----GTGACCGTGAATGCTATCGGGCGCTACCG 200  
 DB 187 GGGGCTGGGGCGGGCGGCAATGCGTGTCCGTGACAGTGTGCTGAAGGCGCGGGCGG 246  
 OY 201 GGAATGCGGACCAACCACTTGTACTGTGGAGCATGGCCGCTGTCCGACCTACTAT 260  
 DB 247 GCGCGCGGGCGCTGCGCCACCACTGTCTAGCTGCGGCTGCGGGCGCTGCTGCTCT 306  
 OY 261 CTGCTGCGGGGTGCGCTGCTGCACTGTACCGGCTC---TGGCGCTGCGGGCGCTGAGTGT 317  
 DB 307 GCTGTGCGGGGTGCGGGTGAAGCTTCTACAGCTTCTGTGTGTTCCATACCTCTGGGTCT 366  
 OY 318 CGGCGCGCTGCTGCGGCTGTCTCTTACGTGGGCGAGGGCTGACCTTACGCGCACT 377  
 DB 367 CGGGAACCTGGGCTGCGGGCGGCTACTACTTGTGTGACAGAGTGTGCGCTACGCGCACT 426  
 OY 378 GCTGCAATGACCGCGCTCAAGCTGTAGCGCTACTGTGGCATTTGGCCCGGCTCGCGG 437  
 DB 427 GCTGAGGTGTGCAAGCGCTGAGCGCGGAGCGCTGCTTACCGGTGTGCGAGCGCTCTG 486  
 OY 438 CGGCGCTTGTGTCAACCGGCGCGCGGCTCGCGCGCTCATGCTGTGCTGTGGGCGGTGAC 497  
 DB 487 CCGGAGCTGTGTAGCGCGACCGCGGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 546  
 OY 498 GCTGCTCTGCGCGGCTTCTTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539  
 DB 547 GCTGTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 588

RESULT 12  
 CNS036WH 941 bp DNA linear GSS 15-MAY-2000  
 LOCUS  
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone

216F22 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.  
 AL230570  
 VERSION  
 GSS; genome survey sequence.  
 AL230570.1 GI:7889565  
 KEYWORDS  
 SOURCE  
 ORGANISM

## REFERENCE

AUTHORS

## TITLE

## REFERENCE

AUTHORS

## TITLE

## JOURNAL

REFERENCE

## JOURNAL

AUTHORS

## JOURNAL

COMMENT

## FEATURES

source

## BASE COUNT

ORIGIN

Query Match 6.0%; Score 122.8; DB 17; Length 941;  
 Best Local Similarity 53.4%; Pred. No. 1.2e-16;  
 Matches 318; Conservative 2; Mismatches 270; Indels 5; Gaps 3;

OY 24 CGACGCCCCGAGGGGGCGCGGAGCCCGGCGCGCGCGCTTGCAGACGAGCG 83  
 DB 605 CGAGCGCAACCGCGCGGAGCTGTAGAGCGGCTGTGAGCGAGAGGCGTACTGCGCAG 546  
 OY 84 CCGGTGCTGCGCCCTTCCCTGCGGGCGCTGTGCGGCTGACCGCTGTGCTGCT 143  
 DB 545 GTACCGGGGCGCGCGGAGTCCCGCTTCTGCGCGCTGTGCGGCTGCGCAT 486  
 OY 144 GTTGTGTGTGCGGGGTGAGCGGCAAGCTGTGACCTGTGATGTGGCGCTTACCGGGA 203  
 DB 485 CTTCCTGCGGGGTGCGGCGGCACTCCCTGACCTGCGCGCTGATCTCGCTTCCGGGCG 426  
 OY 204 CATGCGAACCAACCACTTGTACTGTGGCAGCAAGCGCGTGTCCGACCTAC---TCAT 260  
 DB 425 GANCGCAAGCCACCACTTACTTGTGTGA-CTGTGGCGGCTGCGACCTGTGCTGT 367  
 OY 261 CCGTGTGCGGGCTGCGGCTTGTGACCTGTGACCTGTGCGGCTGCGGCGCTTGGGTGTG 320  
 DB 366 GCTGTGGGGAATGCGCTGTGAGCTGTGACACTGTGTGCGGAATTAACCTTCTGCTG 307  
 OY 321 GCGGCTGTCTGCGCGGCTGCTTCTGTAGTGTGGCGAGAGGCTGACCTTACCGCACTGCT 380  
 DB 306 GGAAGGGGGCTGTACTTCCGACCTTCTGTTCCAG-ACGATGTATTCCTTCATCTCT 248  
 OY 381 GCAATGACCGGCTGAGGCTGAGGCTTACTGTGGCTATGTGCGCGCGCGCTCGCGGCGG 440







2/0 GCAGCCGTCGACCCGCTCCGGCCCTGGGTGTTCCGGCCGCTGCT 321

Db 270 CTTGCCCCCTGAGACTCTATAGATGTGACACACTACCTCTCTCTGCTGAGGCTGTG 329  
Qy 330 CTGCGCCCTGTCCCTCTACGTGGGCGAGGCTGACCTAGCCAGCTGCTGACATGAC 389  
Db 330 CTGCTATTTCGACGACCTACTGTTTGTAGATGTCTGCTGCGCTCACTGTCTCAACGTAC 389  
Qy 390 CCGCTCAGGTGTGAGGCTGACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCTGCT 449  
Db 390 TGCCCTGAGGTGAACGCTATGTGGCGCTGTGTCACCTCAGGCTCAGGCTCATGTGT 449  
Qy 450 CACCCGCGCGCGCTGCGCGCGCTCATGCTGTGCTGTGCGCGCTGCTGCTGCTG 509  
Db 450 GACGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCT 509  
Qy 510 CCGTCCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568  
Db 510 CCGTCCCAACACGAGCTGACGCGCATCCGCGCATGCTGCTGCTGCTGCTGCTGCT 568

## RESULT 2

US-10-083-168-82  
Sequence 82, Application US/10083168  
Publication No. US20030023069A1  
GENERAL INFORMATION:  
APPLICANT: Liaw, Chen W.  
APPLICANT: Chalmers, Derek T.  
APPLICANT: Behan, Dominic P.  
APPLICANT: Maciejewski-Lenior, Dominique  
APPLICANT: Leonard, James N.  
APPLICANT: Ortuno, Daniel  
APPLICANT: Lin, I-Lin  
TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitutively Act  
FILE REFERENCE: AREN-0320  
CURRENT APPLICATION NUMBER: US/10/083.168  
CURRENT FILING DATE: 2002-02-26  
NUMBER OF SEQ ID NOS: 102  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 82  
LENGTH: 1212  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: No. US20030023069A1 Sequence  
US-10-083-168-82

Query Match 5.8%; Score 119; DB 9; Length 1212;  
Best Local Similarity 54.5%; Pred. No. 3.5e-20;  
Matches 261; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

Qy 93 GCGCTTCCCTGGGGGGCTGGGCGGAGACCGCTGTGCGCTGCTGCTGCTGCTGCT 152  
Db 90 GCGCCAGACAGACAGCTGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149  
Qy 153 GCGGCTGAGCGGACAGCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212  
Db 150 GGGGCTGTGGGAGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209  
Qy 213 CACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269  
Db 210 GCGTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269  
Qy 270 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329  
Db 270 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329  
Qy 330 CTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389  
Db 330 CTGCTATTTCGACGACCTACTGTTTGTAGATGTCTGCTGCGCTCACTGTCTCA 389  
Qy 390 CCGCTCAGGTGTGAGGCTGACCTGCGCATCTGCGCGCGCTGCGCGCGCTGCTGCT 449

Db 390 TGCCCTGAGGTGAACGCTATGTGGCGCTGTGTCACCTCAGGCTCAGGCTCATGT 449  
Qy 450 CACCCGCGCGCGCTGCGCGCGCTCATGCTGTGCTGTGCGCGCTGCTGCTGCTGCT 509  
Db 450 GACGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCT 509  
Qy 510 CCGTCCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568  
Db 510 CCGTCCCAACACGAGCTGACGCGCATCCGCGCATGCTGCTGCTGCTGCTGCTGCT 568

## RESULT 3

US-09-104-063-5  
Sequence 5, Application US/09104063  
Patent No. US20020168356A1  
GENERAL INFORMATION:  
APPLICANT: Lee, James  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PFAA Receptors  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/104.063  
FILING DATE: 24-June-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/701265  
FILING DATE: 22-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/664228  
FILING DATE: 06-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/076093  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/810782  
FILING DATE: 19-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Love, Richard B.  
REGISTRATION NUMBER: 34,659  
REFERENCE/DOCKET NUMBER: P0706P2C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-5530  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1679 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
US-09-104-063-5

Query Match 5.1%; Score 103.6; DB 9; Length 1679;  
Best Local Similarity 48.7%; Pred. No. 3.1e-16;  
Matches 348; Conservative 0; Mismatches 354; Indels 12; Gaps 2;

Qy 115 GTGCGGAGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174  
Db 522 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581  
Qy 175 ACCGTGATGCTGATGCGGGGCTGACCGGAGATGCGGACCAACCAACTGTGACTGCTG 234

```

Db 582 GTCTGTGATCTCTGAGGCGACCGGCGAGACGAGTTCACGGAGACCTTCCGTTC 641
Qy 235 AGCATGCGCGCTGTCGACCTACTCATCTGCTCGGGCTGCGCTGACCTGACCGCTC 294
Db 642 CACCTGGCGGTGCGGACCTCTGCTGCTTCTTATTTGCTTCCCTTCCGCTG-----GCC 635
Qy 295 TGGCGCTGCGGCGCTGAGTGTTCGGGCGGCTGCTGCGCGCTGCTTCCCTTCACTGAGG 354
Db 696 GAGGCGCTGAGGCTGAGGCTGAGGAGCTTCTCTGCAAACTGATGATGCGCTGAC 755
Qy 355 GAGGCGCTGAGGCTGAGGCTGAGGAGCTGAGGAGCTGAGGCTGAGGCTGAGGCTGAGG 414
Db 756 AAGTCACTTCTACTGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
Qy 415 GCCATGTCGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
Db 816 GCCATGTCGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
Qy 475 ATGCTGTGCTGTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
Db 876 TGTGAGGACATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
Qy 535 GTGAGGAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
Db 936 GTGAGGAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
Qy 595 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
Db 996 GCAAGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1055
Qy 649 TCCCGCGCTGCGGCGCGGAGACCGGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 708
Db 1056 CCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1115
Qy 709 CCGAGCGCGCGGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
Db 1116 CCGAGCGCGCGGAGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1175
Qy 769 TTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
Db 1176 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229

RESULT 4
US-09-823-114-18
; Sequence 18, Application US/09823114
; Patent No. US20020061554A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, CHRISTOPHER J.
; KEITH, DUANE E.
; TITLE OF INVENTION: OPIOID RECEPTOR GENES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & ROEBSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/823,114
; FILING DATE: 29-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/148,351
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20526, 22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSNPOERSWSH
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1805 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..1119
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-823-114-18

```

```

Query Match 5.0%; Score 102; DB 10; Length 1805;
Best Local Similarity 51.0%; Pred. No. 8.3e-16;
Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

Qy 4 GGCAGCCCTTGAAACGGACGACGAGCCCGAGGAGGCGCGGAGCGCGCGCGCGCGCG 63
Db 52 GGCAGGACACCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 111
Qy 64 CTGCGCGCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 123
Db 112 CATCTGTGCTCAATGACGACGACGAGGCGCTTCTGCGCGCTGCGCGCTGCGCGCTG 171
Qy 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
Db 172 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 231
Qy 184 CTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 243
Db 232 GTATATCTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 291
Qy 244 GTGTCGAGCTACTCATATCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 303
Db 292 CTGCGGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 348
Qy 304 CGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 363
Db 349 GCGCTTCTGCGCGCTTGGAGATGCGCTGTGCAAGACATGCTATGACTATGACTACAC 408
Qy 364 ACTTACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Db 409 ATGTTACGACGACCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTGC 468
Qy 424 CGCGCGCTGCGGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
Db 469 CACCCATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
Qy 484 CTCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 529
Db 529 ATCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 574

RESULT 5
US-09-905-186A-9
; Sequence 9, Application US/0905186A
; Publication No. US20030008289A1
; GENERAL INFORMATION:
; APPLICANT: Kreek, Mary Jeanne
; APPLICANT: Laforge, Karl Steven
; TITLE OF INVENTION: Alleles of the Human Orphanin
; TITLE OF INVENTION: FQ/No. US20030008289A1; Orphanin Receptor Gene, Diagnostic Methods
; FILE REFERENCE: 600-1-284N
; CURRENT APPLICATION NUMBER: US/09/905,186A
; CURRENT FILING DATE: 2001-10-19

```

PRIOR APPLICATION NUMBER: US 60/218,205  
 PRIOR FILING DATE: 2000-07-14  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 9  
 LENGTH: 1829  
 TYPE: DNA  
 ORGANISM: homo sapiens  
 US-09-905-186A-9

Query Match  
 Best Local Similarity 51.0%; Score 102; DB 9; Length 1829;  
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 63  
 DB 76 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 135  
 QY 64 CTGCGCCCTTGCGAGACGCGCGCGCTGCGCGCTTCCCTGAGGCCCAACAGTCTGCCCCG 123  
 DB 136 CATGTGCTCAATGCCAGCAGCGGCTTCTGCCCCCTGAGGCTCAAGTCAATC 195  
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183  
 DB 196 GTGGGGCTTACCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255  
 QY 184 CTGATCGGGGCTTACCGGAGCATCGGACACCACTTGTACTTGGGAGCATGCGC 243  
 DB 256 GTATCTCTCAGGACACCAAAATGAAAGACACCAATATTATTTAACTTTAGCTG 315  
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 303  
 DB 316 CTGCGCAGCACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372  
 QY 304 CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
 DB 373 GCGCTTGGCGCTTGGGAATGCGCTGCAAGACAGTCAATGACATGACTAACAAC 432  
 QY 364 ACCTACGCGCAGCTGCTGCAATGACGCGCTGCAAGCTGCAAGCTGCAAGCTG 423  
 DB 433 ATGTTCACGACGACCTTACCTTACCTGCAATGATGATGATGATGATGATGATG 492  
 QY 424 CGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
 DB 493 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552  
 QY 484 CTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529  
 DB 553 ATCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

RESULT 6  
 US-09-905-186A-10  
 Sequence 10, Application US/09905186A  
 Publication No. US2003008289A1  
 GENERAL INFORMATION:  
 APPLICANT: Kreek, Mary Jeanne  
 APPLICANT: LaForge, Karl Steven  
 TITLE OF INVENTION: Alleles of the Human Orphanin  
 TITLE OF INVENTION: FO/No. US2003008289A1: Orphanin Receptor Gene, Diagnostic Methods  
 TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon  
 FILE REFERENCE: 600-1-284N  
 CURRENT APPLICATION NUMBER: US/09/905,186A  
 PRIOR FILING DATE: 2001-10-19  
 PRIOR APPLICATION NUMBER: US 60/218,205  
 PRIOR FILING DATE: 2000-07-14  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 10  
 LENGTH: 1829  
 TYPE: DNA  
 ORGANISM: homo sapiens  
 US-09-905-186A-10

Query Match  
 Best Local Similarity 51.0%; Score 102; DB 9; Length 1829;  
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 63  
 DB 76 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 135  
 QY 64 CTGCGCCCTTGCGAGACGCGCGCGCTGCGCGCTTCCCTGAGGCCCAACAGTCTGCCCCG 123  
 DB 136 CATGTGCTCAATGCCAGCAGCGGCTTCTGCCCCCTGAGGCTCAAGTCAATC 195  
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183  
 DB 196 GTGGGGCTTACCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255  
 QY 184 CTGATCGGGGCTTACCGGAGCATCGGACACCACTTGTACTTGGGAGCATGCGC 243  
 DB 256 GTATCTCTCAGGACACCAAAATGAAAGACACCAATATTATTTAACTTTAGCTG 315  
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 303  
 DB 316 CTGCGCAGCACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372  
 QY 304 CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
 DB 373 GCGCTTGGCGCTTGGGAATGCGCTGCAAGACAGTCAATGACATGACTAACAAC 432  
 QY 364 ACCTACGCGCAGCTGCTGCAATGACGCGCTGCAAGCTGCAAGCTGCAAGCTG 423  
 DB 433 ATGTTCACGACGACCTTACCTTACCTGCAATGATGATGATGATGATGATGATG 492  
 QY 424 CGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
 DB 493 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552  
 QY 484 CTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529  
 DB 553 ATCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

RESULT 7  
 US-09-905-186A-11  
 Sequence 11, Application US/09905186A  
 Publication No. US2003008289A1  
 GENERAL INFORMATION:  
 APPLICANT: Kreek, Mary Jeanne  
 APPLICANT: LaForge, Karl Steven  
 TITLE OF INVENTION: Alleles of the Human Orphanin  
 TITLE OF INVENTION: FO/No. US2003008289A1: Orphanin Receptor Gene, Diagnostic Methods  
 TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon  
 FILE REFERENCE: 600-1-284N  
 CURRENT APPLICATION NUMBER: US/09/905,186A  
 PRIOR FILING DATE: 2001-10-19  
 PRIOR APPLICATION NUMBER: US 60/218,205  
 PRIOR FILING DATE: 2000-07-14  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 11  
 LENGTH: 1829  
 TYPE: DNA  
 ORGANISM: homo sapiens  
 US-09-905-186A-11

Query Match  
 Best Local Similarity 51.0%; Score 102; DB 9; Length 1829;  
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 63  
 DB 76 GGCAGCCCTTGAAACGGCAGACGCGCCCGAGGGGCGGGAGCCGCGCTGCGCGG 135

QY 64 CTGCCGCTTGAGAGAGGCGGCTGCTGCTTTCCTTGGGGGCGCTGCTGGTG 123  
 Db 136 CATCTGCTCTAATGCAAGCAGCGCGCTTCTGCTCCCTCGGGCTCAAGGTCAACATC 195  
 QY 124 ACCGCTGTGCTGCTGCTGCTTCTGCTGCTGAGGAGGCAAGTGATCGGATG 183  
 Db 196 GTGGGGCTCTACCTGCGCTGTGTGTGCGAAGGCTCTGAGAACTGCTTGTATATAC 255  
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGACACCACTTGTACTGTGAGCATGAGCC 243  
 Db 256 GTCAATCTCAGGACACCAAAATGAAGACAGCACCATAATATTATTAATCTTAACTGGCC 315  
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCACTGTACCGCTCTGGGCTTG 303  
 Db 316 CTGGCGGACACTGTGCTGCTGCTGAGCGCTT---CAAGGACAGCAATCTCTCTG 372  
 QY 304 CGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
 Db 373 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATGCTTCACTAATCAAC 432  
 QY 364 ACCTAAGCGACGCTGCTGCAATGACCGGCTGACGCTGAGCGCTACCTGAGCATCTGC 423  
 Db 433 ATGTTCACACACACTTCACTCACTGATGATGATGATGATGATGATGATGATGATG 492  
 QY 424 CGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
 Db 493 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552  
 QY 484 CTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529  
 Db 553 ATCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

## RESULT 8

US-10-087-345A-22  
 ; Sequence 22, Application US/10087345A  
 ; Publication No. US20030045696A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Owyang, Chung  
 ; TITLE OF INVENTION: ORPHANIN PQ RECEPTOR NUCLEIC ACIDS  
 ; FILE REFERENCE: UM-06962  
 ; CURRENT APPLICATION NUMBER: US/10/087,345A  
 ; CURRENT FILING DATE: 2002-03-01  
 ; NUMBER OF SEQ ID NOS: 36  
 ; SOFTWARE: PatentIn Version 3.0  
 ; SEQ ID NO 22  
 ; LENGTH: 2534  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-087-345A-22

Query Match 5.0%; Score 102; DB 9; Length 2534;  
 Best Local Similarity 51.0%; Pred. No. 9.9e-16;  
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGGAGCCCTTGAGAGAGGCGGCTGCTGCTTTCCTTGGGGGCGCTGCTGGTG 63  
 Db 160 GGAAGCCACTTCAAGGCAACCTGTCTCTGAGCCCAACCAAGTCTGTGCTGCTG 219  
 QY 64 CTGCCGCTTGAGAGAGGCGGCTGCTGCTTTCCTTGGGGGCGCTGCTGGTG 123  
 Db 220 CATCTGCTCTAATGCAAGCAGCGCGCTTCTGCTCCCTCGGGCTCAAGGTCAACATC 279  
 QY 124 ACCGCTGTGCTGCTGCTGCTTCTGCTGCTGAGGAGGCAAGTGATCGGATG 183  
 Db 280 GTGGGGCTCTACCTGCGCTGTGTGTGCGAAGGCTCTGAGGAACTGCTTGTATATAC 339  
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGACACCACTTGTACTGTGAGCATGAGCC 243  
 Db 340 GTATCTCTCAGGACACCAAAATGAAGACAGCACCATAATATTATTAATCTTAACTGGCC 399  
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCACTGTACCGCTCTGGGCTTG 303

Db 400 CTGGCCGACACTGTGCTCTGACGCTGCTT---CAAGGACAGACATCTCTG 456  
 QY 304 CGGCGCTGAGGCTTGGAGCGGCTGCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTG 363  
 Db 457 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATGCTTCACTAATCAAC 516  
 QY 364 ACCTAAGCGACGCTGCTGCAATGACCGGCTGACGCTGAGCGCTACCTGAGCATCTGC 423  
 Db 517 ATGTTCACGACACCTTCACTCACTGATGATGATGATGATGATGATGATGATGATG 576  
 QY 424 CGCGCGCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483  
 Db 577 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636  
 QY 484 CTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529  
 Db 637 ATCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682

## RESULT 9

US-09-905-186A-1  
 ; Sequence 1, Application US/09905186A  
 ; Publication No. US20030008289A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kreek, Mary Jeanne  
 ; APPLICANT: LaForge, Karl Steven  
 ; TITLE OF INVENTION: Alleles of the Human Orphanin  
 ; TITLE OF INVENTION: FQ/No. US20030008289A1; Orphanin Receptor Gene, Diagnostic Methods  
 ; FILE REFERENCE: 600-1-284N  
 ; CURRENT APPLICATION NUMBER: US/09/905,186A  
 ; PRIOR FILING DATE: 2001-10-19  
 ; PRIOR APPLICATION NUMBER: US 60/218,205  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2602  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-09-905-186A-1

Query Match 5.0%; Score 102; DB 9; Length 2602;  
 Best Local Similarity 51.0%; Pred. No. 1e-15;  
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGGAGCCCTTGAGAGAGGCGGCTGCTGCTTTCCTTGGGGGCGCTGCTGGTG 63  
 Db 731 GGAAGCCACTTCAAGGCAACCTGTCTCTGAGCCCAACCAAGTCTGTGCTGCTG 790  
 QY 64 CTGCCGCTTGAGAGAGGCGGCTGCTGCTTTCCTTGGGGGCGCTGCTGGTG 123  
 Db 791 CATCTGCTCTAATGCAAGCAGCGCGCTTCTGCTCCCTCGGGCTCAAGGTCAACATC 850  
 QY 124 ACCGCTGTGCTGCTGCTGCTTCTGCTGCTGAGGAGGCAAGTGATCGGATG 183  
 Db 851 GTGGGGCTCTACCTGCGCTGTGTGTGCGAAGGCTCTGAGGAACTGCTTGTATATAC 910  
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGACACCACTTGTACTGTGAGCATGAGCC 243  
 Db 911 GTCAATCTCAGGACACCAAAATGAAGACAGCACCATAATATTATTAATCTTAACTGGCC 970  
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCTGCGGCTGCACTGTACCGCTCTGGGCTTG 303  
 Db 971 CTGGCGGACACTGTGCTGCTGCTGAGCGCTT---CAAGGACAGCAATCTCTG 1027  
 QY 304 CGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
 Db 1028 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATGCTTCACTAATCAAC 1087  
 QY 364 ACCTAAGCGACGCTGCTGCAATGACCGGCTGACGCTGAGCGCTACCTGAGCATCTGC 423  
 Db 1088 ATGTTCACGACACCTTCACTCACTGATGATGATGATGATGATGATGATGATGATG 1147





```

RESULT 13
US-09-967-768A-296
; Sequence 296, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689230-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28

```

	Query Match	Best Local Similarity	Score	DB	Length
	52.4%	52.4%	92.6	9	1773
				Pred. No.	1,9e-13
	Matches 228; Conservative	0; Mismatches 204; Indels			Gaps 1;
QY	95	CCCTTCCCGGGGGGCGCCTGATGCGCGGTGAACGGCTGTGTGCGCTGTGCTGTTCGTCGCG	154		
b	361	CCCTGATCCCTTCGCGCCTTGCGGCATCGCCATCTACCGGGGCTCTACTCGGCGCGTGTGCGCCCGGG	420		



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2003, 01:14:09 ; Search time 84 Seconds

(without alignments)  
7447.864 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040  
Sequence: 1 atgggcgcgcgcctggaacg.....acgtgaagacgatggatata 2040

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	13.9	283	4	US-08-993-088A-4
2	283	13.9	283	4	US-08-993-424B-4
3	234.8	11.5	1063	4	US-09-077-675A-1
4	230	11.3	1029	4	US-09-077-675A-4
5	228.8	11.2	1122	4	US-09-077-675A-9
6	228.4	11.2	1092	4	US-09-077-675A-15
7	228.4	11.2	3129	4	US-09-077-675A-14
8	227.8	11.2	1088	4	US-09-077-675A-6
9	153	7.5	836	4	US-09-077-675A-11
10	134.8	6.6	1248	4	US-09-545-944-1
11	127	6.2	1575	3	US-08-858-876A-1
12	127	6.2	1575	3	US-09-472-880-1
13	125.4	6.1	1342	3	US-08-833-399-1
14	125.4	6.1	1342	3	US-09-372-498-1
15	122.2	6.0	1529	3	US-08-858-876A-3
16	122.2	6.0	1529	3	US-09-472-880-3
17	120	5.9	1535	4	US-09-668-680-12
18	115.4	5.7	1176	4	US-09-200-090-3
19	115.4	5.7	1233	4	US-09-200-090-1
20	109.8	5.4	1164	4	US-08-993-088A-6
21	109.8	5.4	1164	4	US-08-993-424B-6
22	106.8	5.2	1116	4	US-08-993-088A-18
23	106.8	5.2	1116	4	US-08-993-088A-19
24	106.8	5.2	1116	4	US-08-993-424B-18
25	106.8	5.2	1119	2	US-08-626-685A-7
26	106.8	5.2	1690	4	US-08-665-034A-1
27	106.6	5.2	1610	1	US-08-056-051-5

28	106.6	5.2	1610	1	US-07-928-611-21	Sequence 21, Appl
29	106.6	5.2	1610	2	US-08-487-811A-21	Sequence 21, Appl
30	106.6	5.2	1610	4	US-09-060-694-21	Sequence 21, Appl
31	106.6	5.2	1610	4	US-09-378-074-21	Sequence 21, Appl
32	106.6	5.2	1610	5	PCT-US93-07370-21	Sequence 21, Appl
33	103.6	5.1	1679	1	US-08-202-056-6	Sequence 6, Appl
34	103.6	5.1	1679	1	US-08-076-093A-5	Sequence 5, Appl
35	103.6	5.1	1679	1	US-08-701-265-5	Sequence 5, Appl
36	103.6	5.1	1679	2	US-08-284-586-5	Sequence 5, Appl
37	103.6	5.1	1679	2	US-08-805-478-5	Sequence 5, Appl
38	103.6	5.1	1679	2	US-08-802-627A-5	Sequence 5, Appl
39	103.6	5.1	1679	2	US-08-801-238-5	Sequence 5, Appl
40	103.6	5.1	1679	2	US-08-801-228-5	Sequence 5, Appl
41	103.6	5.1	1679	3	US-09-104-296-5	Sequence 5, Appl
42	103.6	5.1	1679	5	PCT-US94-06380-3	Sequence 3, Appl
43	103.6	5.1	2818	3	US-08-982-493-7	Sequence 7, Appl
44	103.6	5.1	2818	4	US-08-628-655-1	Sequence 1, Appl
45	102	5.0	1805	4	US-08-405-271A-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-08-993-088A-4  
Sequence 4, Application US/08993088A  
Patent No. 6287855  
GENERAL INFORMATION:  
APPLICANT: Tan, Carina  
TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND  
TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,088A  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/033,851  
FILING DATE: 27-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Heber, Sheldon O.  
REGISTRATION NUMBER: 38,179  
REFERENCE/DOCKET NUMBER: 19846  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1958  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 283 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
MOLECULE TYPE: Other  
NAME/KEY: Other  
LOCATION: 1...283  
OTHER INFORMATION: cDNA probe  
US-08-993-088A-4

Query Match 13.9%; Score 283; DB 4; Length 283;  
Best Local Similarity 100.0%; Pred. No. 3.6e-54;  
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCACTTGTACTGGGAGCATGCGGTGTCCGACCTACTCATCTTCC 265  
DB 1 TGGGACCAACCACTTGTACTGGGAGCATGCGGTGTCCGACCTACTCATCTTCC 60  
QY 266 TCGGCTGCGCTGTGACCTGTACCGCTCTGCGGCGCCCTGGGGTGTGGGGCGC 325  
DB 61 TCGGCTGCGCTGTGACCTGTACCGCTCTGCGGCGCCCTGGGGTGTGGGGCGC 120  
QY 326 TGTCTGCGCGCTGTGACCTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 385  
DB 121 TGTCTGCGCGCTGTGACCTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 386 TGACCGCGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 445  
DB 181 TGACCGCGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 240  
QY 446 TGTGACCGCGCGCGCGCGCTGCGGCGGCTGATGCTGTGCTG 488  
DB 241 TGTGACCGCGCGCGCGCGCTGCGGCGGCTGATGCTGTGCTG 283

## RESULT 2

US-08-993-424B-4  
Sequence 4, Application US/08993424B  
Patent No. 6337206  
GENERAL INFORMATION:  
APPLICANT: Tan, Carina  
APPLICANT: Kolakowski, Lee F., Jr.  
TITLE OF INVENTION: MOUSE GLANIN RECEPTOR GALR2 AND  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,424B  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/033,851  
FILING DATE: 27-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Heber, Sheldon O.  
REGISTRATION NUMBER: 38,119  
REFERENCE/DOCKET NUMBER: 19846NP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1958  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 283 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other  
FEATURE:  
NAME/KEY: Other  
LOCATION: 1...283  
OTHER INFORMATION: cDNA probe

US-08-993-424B-4  
Query Match 13.9%; Score 283; DB 4; Length 283;  
Best Local Similarity 100.0%; Pred. No. 3.6e-54;  
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCACTTGTACTGGGAGCATGCGGTGTCCGACCTACTCATCTTCC 265  
DB 1 TGGGACCAACCACTTGTACTGGGAGCATGCGGTGTCCGACCTACTCATCTTCC 60  
QY 266 TCGGCTGCGCTGTGACCTGTACCGCTCTGCGGCGCCCTGGGGTGTGGGGCGC 325  
DB 61 TCGGCTGCGCTGTGACCTGTACCGCTCTGCGGCGCCCTGGGGTGTGGGGCGC 120  
QY 326 TGTCTGCGCGCTGTGACCTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 385  
DB 121 TGTCTGCGCGCTGTGACCTGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 386 TGACCGCGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 445  
DB 181 TGACCGCGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 240  
QY 446 TGTGACCGCGCGCGCGCGCTGCGGCGGCTGATGCTGTGCTG 488  
DB 241 TGTGACCGCGCGCGCGCGCTGCGGCGGCTGATGCTGTGCTG 283

## RESULT 3

US-09-077-675A-1  
Sequence 1, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pai, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Fong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452  
REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1063 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

US-09-077-675A-1

Query Match 11.5%; Score 234.8; DB 4; Length 1063;  
Best Local Similarity 69.3%; Pred. No. 2.8e-43;  
Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY CGCCCTTCCCTGGGGGCGCTGCTGCGGTACCGCTGTGTGCTGTGCTGTG 151  
DB CGCTCTTCCCAAGCGCGCTGTGGGGGCGTACCGCCACCTGCGGTGCTGTG 128  
QY TCGGGGTGAGCGGCAACGCTGATCGGTATGCTGATGCGGGCTACCGGGACATGCGGA 211  
DB TGGGTATGCGGGGCACTGCTCAAGATCTGTGTATGATGATGCTTCCCGAGATGCGCA 188  
QY CCACCAACCACTTGTACTGTGGGAGACATGCGGTGCTGACCTACTCATCTGTGCGGC 271  
DB CCACCAACCACTTGTACTGTGACATGCGGTGCTGACCTACTCATCTGTGCGGC 248  
QY 272 TGGCGTTGACTGTATCGGCTCTGTGGCGCTCGCGGCTTGGGGTTCGGGCGGCTGCT 331  
DB 249 TGGCGTTGACTGTATCGGCTCTGTGGCGCTCGCGGCTTGGGGTTCGGGCGGCTGCT 308  
QY 332 GCGCGCTGCTCCTCTACGAGGGGCGAGGGGCTGACCTAGCGCAAGCTGTGACATGACCG 391  
DB 309 GCAAACTTTCAGTCTGTATGCGAGAGCTGCACTACCGCAAGTGTGACATGACCG 368  
QY 392 CGCTAGCGTGAAGCGCTACCTGAGCATCTGCGCGCTCGCGCGCGCTTGTGTGA 451  
DB 369 CGTGAAGGCTGAGAGGCTACTTGGCATCTGCTTCCCGCTGGGGGCGAAGTATGTGTA 428  
QY 452 CCGCGCGCGCGCTCGCGCGCTCATGCTGTGCTGTGGGCGGTGCGCTGTGCTGCGG 511  
DB 429 CCAAGGGCGGGGTAAAGCTGTGATCTGTGATCTGTGGGCGGTGCGCTGTGCGGCGG 488  
QY 512 GTCCCTTCTGTCTGTGAGGGGCTGAGCAGGACCGCGGCA 553  
DB 489 GCGCCATCTTCTGTGCTGTGAGGAGTGAAGATGATTAACGGCA 530

## RESULT 4

US-09-077-675A-4  
Sequence 4, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pal, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,675A  
FILING DATE: 3-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cocuzzo, Anna L.  
REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1273  
TELEFAX: 732-594-4720  
TELEX:

INFORMATION FOR SEQ. ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1029 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-077-675A-4

Query Match 11.3%; Score 230; DB 4; Length 1029;  
Best Local Similarity 68.6%; Pred. No. 3.2e-42;  
Matches 317; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 92 CGCCCTTCCCTGGGGGCGCTGCTGCGGTACCGCTGTGTGCTGTGCTGTG 151  
DB 266 CGCTCTTCCCAAGCGCGCTGTGGGGGCGTACCGCCACCTGCGGTGCTGTG 325  
QY 152 TCGGGGTGAGCGGCAACGCTGATCGGTATGCTGATGCGGGCGCTTACCGGACATGCGGA 211  
DB 326 TGGGTATGCGGGGCACTGTCTCAAGATCTGTGTATGATGATGCTTCCCGAGATGCGCA 385  
QY 212 CCAACCAACCACTTGTACTGTGGGAGACATGAGCGGTGCTGACCTACTCATCTGTGCGGC 271  
DB 386 CCAACCAACCACTTGTACTGTGACATGAGCGGTGCTGACCTACTCATCTGTGCGGC 445  
QY 272 TGGCGTTGACTGTATCGGCTCTGTGGCGCTCGCGGCTTGGGGTTCGGGCGGCTGCT 331  
DB 446 TGGCGTTGACTGTATCGGCTCTGTGGCGCTCGCGGCTTGGGGTTCGGGCGGCTGCT 505  
QY 332 GCGCGCTGCTCCTCTACGAGGGGCGAGGGGCTGACCTAGCGCAAGCTGTGACATGACCG 391  
DB 506 GCAAACTTTCAGTCTGTATGCGAGAGCTGCACTACCGCAAGTGTGACATGACCG 565  
QY 392 CGCTAGCGTGAAGCGCTACCTGAGCATCTGCGCGCTCGCGCGCGCTTGTGTGA 451  
DB 566 CGTGAAGGCTGAGAGGCTACTTGGCATCTGCTTCCCGCTGGGGGCGAAGTATGTGTA 625  
QY 452 CCGCGCGCGCGCTCGCGCGCTCATGCTGTGCTGTGGGCGGTGCGCTGTGCTGCGG 511  
DB 626 CCAAGGGCGGGGTAAAGCTGTGATCTGTGATCTGTGGGCGGTGCGCTGTGCGGCGG 685  
QY 512 GTCCCTTCTGTCTGTGAGGGGCTGAGCAGGACCGCGGCA 553  
DB 686 GCGCCATCTTCTGTGCTGTGAGGAGTGAAGATGATTAACGGCA 727

## RESULT 5

US-09-077-675A-9  
Sequence 9, Application US/09077675A  
Patent No. 6242199  
GENERAL INFORMATION:  
APPLICANT: Pal, Lee-Yuh  
APPLICANT: Feighner, Scott C.  
APPLICANT: Howard, Andrew D.  
APPLICANT: Pong, Sheng-Shung  
APPLICANT: Van Der Ploeg, Leonardus H.T.  
TITLE OF INVENTION: RECEPTOR ASSAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible





```

RESULT 10
US-09-545-944-1
: Sequence 1, Application US/09545944
: Patent No. 6461836
: GENERAL INFORMATION:
: APPLICANT: AMES, ROBERT
: APPLICANT: ELSHOURBAY, NABIL
: APPLICANT: MICHALOVICH, DAVID
: APPLICANT: SARAU, HENRY
: APPLICANT: SHABON, USMAN
: APPLICANT: VAMTER, LISA
: TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
: TITLE OF INVENTION: (AXOR34) AND SCREENING METHODS THEREOF
: FILE REFERENCE: GP/0657-1
: CURRENT APPLICATION NUMBER: US/09/545,944
: CURRENT FILING DATE: 2000-04-10
: PRIOR APPLICATION NUMBER: US 09/435,384
: PRIOR FILING DATE: 1999-11-05
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PasteSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1248
: TYPE: DNA

```





ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC  
STREET: 400 Seventh Street  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/472,880  
FILING DATE: 28-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR 9723204  
FILING DATE: 17-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,049  
INFORMATION FOR SEQ ID NO: 1  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1575 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 37..1266  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-472-880-1

Query Match 6.2%; Score 127; DB 4; Length 1575;  
Best Local Similarity 55.2%; Pred. No. 2,7e-19;  
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

27 CGGCCCCGAGGGGGGCGGGAGCCCGCTGCGCCGCTGCGCCCTTGCCAGCGCCGCG 86  
45 CAGCAGCCCGCGGCGCCCGCGCCAGCTCCACCAGCGGCGTGAAGCCCGCGCT 104  
87 CTGCTGCGCCCTTCCCTGCGGGGCGCTGCGCGGCTGACCGCTGAGTGGCTGCTT 146  
105 GGGCGTGAACATCGCTCTTGGGCGAAGTGTCTTACCGCGCTTACGCACTCATCTG 164  
147 CGTCTGCGGGGTGAGCGGCAACGTG-----GTGACCGTGAATGCTGATCGGCGCTACCG 200  
165 GGGCGTGGGGCGCGCGGCAATGCGTCCGTGACGCTGCTGAAGCGCGGCGCG 224  
201 GGAATGCGGACCAACCAACTTGTATCTGGGCAAGATGCGCGCTGTCCGACTATCAT 260  
225 GCGCGCGGGGCGCTGCGCCCAACGCTCAGCTGCGCGCTGCGGGCGCTGCTGCTCT 284  
261 CCGTCTGCGGGGCTGCGCTTCAACCTGTACCGCTGCGCGCTCGCGG---CCCTGGGCTTT 317  
285 GCTGTGTGGGGTGGCGGCTGTACAGCTTGTGTGTTCACATACCCCTGGGCTT 344  
318 CGGCGCGCTGCTGCGCGCTGTCCCTTACGTGGGCGAGGAGCTGACCTTACGCGCGCT 377  
345 CGGCGACTGGGCTGCGCGGCTACTTCTGTGACGAGCTGTGCGCTACGCGCAGCT 404  
378 GTTGACATGACCGCGCTCAGCGTGAAGCGCTACTTGGCCATCTGCGCGCGCTCGCGC 437  
405 GCTGAGCGTGGCAGAGCTGAGCGCGCGACGCTGCTTACCGCTGTCCAGCCCTTGTGC 464  
438 CCGGCTTTGTGTCACCGCGGGCGCGCGCGCGCTCATCGCTGTGCTTGGGCGCGTGGC 497  
465 CCGAGAGCTGTGACGCGCAGCGCGGAGCCGGGTGGTGTGCTGTGTGGGCGCGCTC 524  
498 GCTGCTCTTCCGCGCTCTTGTCTTGTCTGTGTGCGCTGAGCAGAACCCCG 550

Db 525 GCTGCGGCTGCGCCCTGCTCCATGCGCCCTCATCATGAGGAGCAGCACTACCTG 577

RESULT 13  
US-08-832-399-1  
Sequence 1, Application US/08832399  
Patent No. 6008050  
GENERAL INFORMATION:  
APPLICANT: Bergsma, Dirk  
APPLICANT: Shabon, Usman  
TITLE OF INVENTION: NOVEL HUMAN NEUROTENSIN RECEPTOR TYPE 2  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: PA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,399  
FILING DATE: 02-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: King, William T  
REGISTRATION NUMBER: 30,954  
REFERENCE/DOCKET NUMBER: GH50020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5515  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-832-399-1

Query Match 6.1%; Score 125.4; DB 3; Length 1342;  
Best Local Similarity 55.0%; Pred. No. 5.8e-19;  
Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

27 CGGCCCCGAGGGGGGCGGGAGCCCGCGTGGCGCGCGCTTGGCAGCGAGCGCG 86  
61 CAGCAGCCCGCGGCGCCCGCGCCAGCTCCACCAGCGGCGTGAAGCGCCCGGCT 120  
87 CTGCTGCGCCCTTCCCTGCGGGGCGCTGCGCGGCTGACCGCTGTGTGCTGCTGTT 146  
121 GGGCGTGAACATCACTCTGTGGCCAAAGTGTCTTCACTGCGCTTACGCACTCATCTG 180  
147 CGTCTGCGGGGTGAGCGGCAACGTG-----GTGACCGTGAATGCTGATGGGCGCTAACCG 200  
181 GCGCGTGGGCGCGGGCGAATGCGCTGCTGCTGCACTGTGTGAAGCGCGGCGCG 240  
201 GGAATGCGGACCAACCAACTTGTATCTGGGCAAGATGCGCGCTGTCCGACTATCAT 260  
241 GCGCGCGGGGCGCTGCGCCACACAGCTGTCACTGCGCGCTGCGGCGCTGCTGCT 300  
261 CCGTCTGCGGCTGCGCTTCACTGTACCGCTTGTGCGCGCTCGCG---CCCTGGGCTTT 317  
301 GCTGTGCGGCGTGGCGGCTGGAACCTTACAGCTTGTGTGTTCACACTACCCCTGGGCTT 360



Db 238 GCGCGTGGCGTACACAGTGTCTAGCGGCTCTGACCCCTGTGCTACTGTGTCTAG 297  
QY 270 GCTGCGGTTGACCTGTACCGCCTCTGGCGCTGGCGG--CCCTGGGGTGTTCGGGCGGCT 326  
Db 298 CATGCCCATGAGAGCTCTACAACTTCGTGTGTCTCCACTACCCATGGGTCTTCGGCGATCT 357  
QY 327 GCTTGGCGCGCTGTCCCTTACGTGAGCGAGGGCTGCACCTACGCCACGCTGTGCACAT 386  
Db 358 GGGCTGCGGCTGTATTACTTCTGTGCGCGAGCTGTGCGCTAGCGCACAGTGTGAGCGT 417  
QY 387 GACCGGCTCAGCGTGAAGGCTTACCTGAGCCATCTGCGCCGCTCCGCGCGCGCTT 446  
Db 418 TGCAGAGCTTAAGGCGAGAGGCTGCTGCGCGGTGTGCTCAGCGCGCTGCGCGCGCGCT 477  
QY 447 GGTACCGCGCGCGCGGCTCCGCGCGCTCATGCTGTGCTGTGGGCGGAGGCTGTCTC 506  
Db 478 TCTCACCGCGCGCGCGACCCGCGCGCTGTGTACTGTGTGTGGGTGCGCTCTGGGCT 537  
QY 507 TGCAGGTTCCCTTCTTGTCTGTGAGCGGTGAGCAGGA 545  
Db 538 TGCCCTGCCCATGGCGGTTATCATGGGACAGAGCAGCA 576

Search completed: March 16, 2003, 10:51:09  
Job time : 88 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 08:50:18 ; Search time 480 Seconds  
(without alignments)  
9570.993 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040  
Sequence: 1 atggcgagccctcggaacg.....acgtgaagcagtgagataa 2040

Scoring table: IDENTITY NUC  
Gapop 10-0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: N\_Geneseq\_101002.\*  
2: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2040	100.0	2040	21	AAA46116	Human G protein co
2	2021.2	99.1	3066	21	AAZ5402	Genomic sequence o
3	1047.2	51.3	1161	22	AAZ53684	Short form of moti
4	1047.2	51.3	1390	21	AAZ5404	cDNA encoding the
5	896.2	43.9	1239	22	AAZ5403	cDNA encoding the
6	896.2	43.9	1239	22	AAZ53683	Long form of moti
7	896.2	43.9	1239	22	AAZ5449	Nucleotide sequenc
8	896.2	43.9	1506	22	AAZ5449	Human GPR38 varian
9	669.8	32.8	1203	22	AAZ5448	Nucleotide sequenc

10	541.2	26.5	813	22	AAZ5447	Nucleotide sequenc
11	501.2	24.6	1179	24	ABQ47146	Oligonucleotide fo
12	501.2	24.6	1179	24	ABQ47147	Oligonucleotide fo
13	437.2	21.4	1179	24	ABQ47148	Oligonucleotide fo
14	437.2	21.4	1179	24	ABQ47149	Oligonucleotide fo
15	283	13.9	283	19	AAZ44930	Galatin receptor G
16	283	13.9	283	19	AAZ32651	Galatin receptor G
17	283	13.9	283	19	AAZ32651	Galatin receptor G
18	283	13.9	283	24	ABK14060	Rat galatin recept
19	248.2	12.2	1050	21	AAZ61492	cDNA encoding canl
20	234.8	11.5	1063	18	AAZ68662	Pig growth hormone
21	234.8	11.5	1063	18	AAZ69754	Swine growth hormo
22	233.2	11.4	1029	18	AAZ68663	Pig growth hormone
23	233.2	11.4	1029	18	AAZ69755	Swine growth hormo
24	231.6	11.4	1095	21	AAZ45993	cDNA encoding the
25	231.6	11.4	4009	21	AAZ45967	DNA encoding the m
26	228.8	11.2	1122	18	AAZ68665	Human growth hormo
27	228.4	11.2	1092	22	AAZ7800	Human growth hormo
28	228.4	11.2	3129	18	AAZ68667	Rat growth hormone
29	228.4	11.2	3129	18	AAZ68667	Rat growth hormone
30	227.8	11.2	1088	18	AAZ68664	Human growth hormo
31	227.8	11.2	1088	18	AAZ69756	Human growth hormo
32	227.8	11.2	1101	21	AAA30643	Human G protein-co
33	227.8	11.2	1101	21	AAZ30732	DNA encoding human
34	227.8	11.2	1101	21	AAZ51463	Human G-protein-co
35	227.8	11.2	1101	21	AAZ51463	Human G-protein-co
36	227.8	11.2	1101	22	AAZ83680	Human growth-hormo
37	227.2	11.1	1122	18	AAZ69757	Human growth hormo
38	226.8	11.1	1092	18	AAZ69760	Rat growth hormone
39	225.2	11.0	1092	21	AAZ45405	DNA encoding the p
40	225.2	11.0	1092	22	AAZ85450	Nucleotide sequenc
41	184.6	9.0	1098	22	AAZ83681	Human G-protein co
42	153	7.5	836	18	AAZ68666	Human growth hormo
43	153	7.5	836	18	AAZ69758	Human growth hormo
44	134.8	6.6	729	22	AAZ85407	Nucleotide sequenc
45	134.8	6.6	801	22	AAZ50977	Human nPCR15 codi

## ALIGNMENTS

RESULT 1	AAA46116	standard; cDNA; 2040 BP.
ID	AAA46116;	
AC	AAA46116;	
DT	22-AUG-2000	(first entry)
DE	Human G protein coupled receptor hGPR38 (V297K)	cDNA SEQ ID NO:129.
KW	Human; G protein coupled receptor; GPCR; transmembrane receptor;	
KW	identification; agonist; screening; therapeutic; pharmaceutical;	
KW	mutant; ss.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
PN	WO200022131-A2.	
XX		
PD	20-APR-2000.	
XX		
PF	13-OCT-1999;	99WO-US24065.
XX		
PR	13-OCT-1998;	98US-0170496.
PR	12-NOV-1998;	98US-0108029.
PR	20-NOV-1998;	98US-0109213.
PR	27-NOV-1998;	98US-0110060.
PR	16-FEB-1999;	99US-0120416.
PR	26-FEB-1999;	99US-0121852.
PR	12-MAR-1999;	99US-0123944.
PR	12-MAR-1999;	99US-0123945.
PR	12-MAR-1999;	99US-0123946.

PR 12-MAR-1999; 99US-0123948.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 28-MAY-1999; 99US-0137567.  
 PR 30-JUN-1999; 99US-0141448.  
 PR 27-AUG-1999; 99US-0151114.  
 PR 03-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156634.  
 XX  
 XX (AREN-) ARENA PHARM INC.  
 XX  
 XX Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT,  
 PI Gore M, Law CM, Lin I, Lowitz K, White C;  
 XX WPI: 2000-317986/27.  
 XX P-PSDB; AAB02854.  
 DR  
 XX Non-endogenous, human G protein-coupled receptors for screening  
 PT receptor, inverse or partial agonists useful as therapeutic agents -  
 XX  
 PS Example 2; Page 166-168; 187pp; English.  
 XX  
 XX The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCRs), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA6017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;

Query Match 100.0%; Score 2040; DB 21; Length 2040;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2040; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 TGCCGCCGCTCCGCGCCGCGCTTGTGTACCCGCGCCGCGCTTCATCGCT 480  
 DB 421 TGCCGCCGCTCCGCGCCGCGCTTGTGTACCCGCGCCGCGCTTCATCGCT 480  
 QY 481 GTGCTTGGGCGGTGGGCTGTCTTCCGCGCTTGTGTACCCGCGCCGCGCTTCATCGCT 540  
 DB 481 GTGCTTGGGCGGTGGGCTGTCTTCCGCGCTTGTGTACCCGCGCCGCGCTTCATCGCT 540  
 QY 541 CAGACCCCGGATCTCCGATGTCGCGGCGCTCAATGAGCACCGCGGATGCGCTTCG 600  
 DB 541 CAGACCCCGGATCTCCGATGTCGCGGCGCTCAATGAGCACCGCGGATGCGCTTCG 600  
 QY 601 CCTCTGCGCTGTGCGCGCTCTCTGCTTTCGCGCGCGCACCGCGCTTCGCGCTG 660  
 DB 601 CCTCTGCGCTGTGCGCGCTCTCTGCTTTCGCGCGCGCACCGCGCTTCGCGCTG 660  
 QY 661 GGGCCCGAGACCGCGAGCGCGCGCGCTGTACCGCGGAATGCGCGAGCCCGCG 720  
 DB 661 GGGCCCGAGACCGCGAGCGCGCGCGCTGTACCGCGGAATGCGCGAGCCCGCG 720  
 QY 721 CAGCTGGGCGGCTGCGCTGTGAGTCAACACCGCGCTTCTTCTGCGCTT 780  
 DB 721 CAGCTGGGCGGCTGCGCTGTGAGTCAACACCGCGCTTCTTCTGCGCTT 780  
 QY 781 CTGTGCTCAGCATCTCTACAGGCTCATCGGCGGAGCTGTGAGACAGCGCGCGCG 840  
 DB 781 CTGTGCTCAGCATCTCTACAGGCTCATCGGCGGAGCTGTGAGACAGCGCGCGCG 840  
 QY 841 CTGCGAGGCGCGCGCGCTTCCGCGGCGGAGAGAGCGCACCGCGAGCCGCGCTG 900  
 DB 841 CTGCGAGGCGCGCGCGCTTCCGCGGCGGAGAGAGCGCACCGCGAGCCGCGCTG 900  
 QY 901 CGTAAGTGAAGCGCGCGCTGTCCAAAGACGCTGCTGAGTCCGCGCGCGCGGAGCC 960  
 DB 901 CGTAAGTGAAGCGCGCGCTGTCCAAAGACGCTGCTGAGTCCGCGCGCGCGGAGCC 960  
 QY 961 GCGCAAGCGTGGTCCCTTCCCTGCTGCTGCGCGAGCTGTGGGCGCGCTTCCAGCTCC 1020  
 DB 961 GCGCAAGCGTGGTCCCTTCCCTGCTGCTGCGCGAGCTGTGGGCGCGCTTCCAGCTCC 1020  
 QY 1021 TTTTCTATTTGATTCAGCGCTTCCAGCGCGCGGATCTCCATCCCGCGAGAAACATG 1080  
 DB 1021 TTTTCTATTTGATTCAGCGCTTCCAGCGCGCGGATCTCCATCCCGCGAGAAACATG 1080  
 QY 1081 TCTGTCCCGCAGAGCTCTGGGAGACCCAGAGCGCTTGTGAGGTGGATCCCGGATC 1140  
 DB 1081 TCTGTCCCGCAGAGCTCTGGGAGACCCAGAGCGCTTGTGAGGTGGATCCCGGATC 1140  
 QY 1141 CGATTCAAGTAACAGAGCTTTTCCAGAGCTCTGAGACCAAGAAAGAGTGTGTA 1200  
 DB 1141 CGATTCAAGTAACAGAGCTTTTCCAGAGCTCTGAGACCAAGAAAGAGTGTGTA 1200  
 QY 1201 TTTCTTAATCCAAACGCTGTGATGACCAATGAGAGTCTTCAAGTCTTGA 1260  
 DB 1201 TTTCTTAATCCAAACGCTGTGATGACCAATGAGAGTCTTCAAGTCTTGA 1260  
 QY 1261 AGACGAGGAGATTCAATTAAGCTAAATTTTATTAATTAATTAATGATGCTGAAGC 1320  
 DB 1261 AGACGAGGAGATTCAATTAAGCTAAATTTTATTAATTAATTAATGATGCTGAAGC 1320  
 QY 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGCAAGCTTGTGAGATTCTT 1380  
 DB 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGCAAGCTTGTGAGATTCTT 1380  
 QY 1381 TTCAACAGAGAAACAAAACTTGTCTCGAAGTGGGTTTGGAGAGGAGCGCGCAAG 1440  
 DB 1381 TTCAACAGAGAAACAAAACTTGTCTCGAAGTGGGTTTGGAGAGGAGCGCGCAAG 1440  
 QY 1441 CGGCTTTGTAGAGAAATGCTCTTCTGTTATGTCCAGCTTGATTAACATATGAG 1500  
 DB 1441 CGGCTTTGTAGAGAAATGCTCTTCTGTTATGTCCAGCTTGATTAACATATGAG 1500  
 QY 1501 AGCTTAATGAGATTAAAGCAAGTATCATGACGCTGACGCTGATTTTCTT 1560

Db	1501	AGCCTACTATGACGGTTTAAAGCAAGATCATGCAGCCCGACGCCCTGGTCATTTTTTTCT	1560
Oy	1561	GGGGTAGAGATCTGCCTAGGTAGAAGTTTCTTAATTATTTTGCTGTACTTGTATT	1620
Db	1561	GGGGTAGAGATCTGCCTAGGTAGAAGTTTCTTAATTATTTTGCTGTACTTGTATT	1620
Oy	1621	GCAGATGTCTTCCTTGCGGGGTGGGGGTTATTTGCTGCCAATGCTTTGTAAATCCC	1680
Db	1621	GCAGATGTCTTCCTTGCGGGGTGGGGGTTATTTGCTGCCAATGCTTTGTAAATCCC	1680
Oy	1681	GGTGCTGTGCTCTTANGTTCAGTGTGTGTGTTCTGGCATTTATAATTGCTGGTTGCC	1740
Db	1681	GGTGTGCTGTCTTATTTGTGAAGTGTGTGTGTTCTGGCATTTATAATTGCTGGTTGCC	1740
Oy	1741	TTCACAGTTGGCAGATTCATTTACAATAACACGGAAGATTGCGGATGATGTACTTCTCT	1800
Db	1741	TTCACAGTTGGCAGATTCATTTACAATAACACGGAAGATTGCGGATGATGTACTTCTCT	1800
Oy	1801	CAGTACTTTAACATCGTCGCTGCAACTTTTCTATCTGAGCCATCTATCAACCACATC	1860
Db	1801	CAGTACTTTAACATCGTCGCTGCAACTTTTCTATCTGAGCCATCTATCAACCACATC	1860
Oy	1861	CTCTACAACTCATTTCAAAGAATCACAGACGGCGCCTTTAACTGCTGCTCGCAAG	1920
Db	1861	CTCTACAACTCATTTCAAAGAATCACAGACGGCGCCTTTAACTGCTGCTCGCAAG	1920
Oy	1921	AAGTCACAGCCGAGAGGCTTCCACAAGACGAGACACTCGCGGGGAAAGTTCAGAGGGAC	1980
Db	1921	AAGTCACAGCCGAGAGGCTTCCACAAGACGAGACACTCGCGGGGAAAGTTCAGAGGGAC	1980
Oy	1981	ACTGAGAGACACGGTGGGCTTACACCGAGACAAAGCGCTAACGTGAAGACGATGGANTAA	2040
Db	1981	ACTGAGAGAGACACGGTGGGCTTACACCGAGACAAAGCGCTAACGTGAAGACGATGGANTAA	2040
RESULT 2			
ID	AAZ45402	standard; DNA; 3066 BP.	
XX	AAZ45402;		
DT	27-MAR-2000	(first entry)	
DE	Genomic sequence of the motilin receptor gene including 5' UTR.		
KW	Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;		
KM	spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;		
KW	functional defect; neurological disorder; scleroderma; colonoscopy;		
KM	paraneoplastic syndrome; radiation induced dysmotility; diabetes;		
KW	infection; stress-related motility disorder; psychogenic disorder;		
KM	gastroesophageal reflux disease; constipation;		
KW	chronic idiopathic pseudo obstruction; acute faecal impaction;		
KM	postoperative ileus; gallstones; infantile colic; diarrhoea;		
KW	irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;		
KX	endoscopy; duodenal intubation; ds.		
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	misc_feature	1929..1931	
FT		/tag= a	
FT		/note= "imperfect donor site"	
FT	intron	1930..2728	
FT		/tag= b	
FT		/note= "intronic sequence"	
FT	misc_feature	2080..2082	
FT		/tag= c	
FT		/note= "perfect donor site"	
FT	misc_feature	2729..2732	
FT		/tag= d	
FT		/note= "perfect splice acceptor site"	
XX			

PN	M090964436-A1.
XX	16-DEC-1999.
PF	08-JUN-1999; 99WO-US12773.
XX	
PR	12-JUN-1998; 98US-0089098.
XX	
PA	(MERI ) MERCK & CO INC.
XX	
PI	Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
P1	Pong S, Smith RG;
XX	
DR	WPI; 2000-105868/09.
XX	
PT	Novel receptor protein for screening compounds used in treating
PS	irritable bowel syndrome, constipation and other gastric conditions -
BS	Example 1; Fig 1; 44pp; English.
XX	
CC	The present sequence represents the genomic sequence of the motilin
CC	receptor gene, including the 5' untranslated region (5' UTR). This gene
CC	encodes a G-protein coupled receptor, and is designated MTL-R1 (also
CC	GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY5145) and
CC	MTL-R1B (see AAY5146). MTL-R1A is a functional seven transmembrane
CC	domain form, and MTL-R1B is a truncated five transmembrane domain. The
CC	MTL-R1 proteins are used to identify agonists and antagonists which can
CC	be used for treating gastric motility disorders, functional defects,
CC	disorders secondary to neurological disorders e.g. scleroderma,
CC	paraneoplastic syndromes radiation induced dysmotility, diabetes,
CC	infections, stress-related motility disorders, psychogenic disorders,
CC	gastroenteritis, gastro-oesophageal reflux disease, constipation,
CC	chronic idiopathic pseudo obstruction, acute faecal impaction,
CC	postoperative ileus, gallstones, infantile collic, irritable bowel
CC	syndrome, non-ulcer dyspepsion, non-cardiac chest pain and diarrhoea.
CC	They can also be used in the preparation for colonoscopy, endoscopy and
CC	duodenal intubation.
SQ	
Sequence	3066 BP; 585 A; 930 C; 926 G; 625 T; 0 other;
Query Match	99.1%; Score 2021.2; DB 21; Length 3066;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2035; Conservative	0; Mismatches 3; Indels 2; Gaps 1.
OY	1 ATGGGACAGCCCCCTGTGAACGGCAGCGACGCCGCCGAGGGGCGCGGAGCCGCGTGAGCCC 60
DB	1029 ATGGGCACACCCTTGGAAACGGCAGCGACGCCGCCCGAGGGGGCCGGGAGCCGCCGTGAGCCC 1088
OY	61 GCGCTGCCGCTTGCGACGAGGCGCGCGTCGTCGCCCTTCCCTGGGGGCGCTGTGTCGG 120
DB	1089 GCGCTGCCGCTTGCGAGCGAGGCGCGCTGTCGCCCTTCCCTGGGGGCGCTGTGTCGG 1148
OY	121 GTGACCGGTGTGTGCGCTTGTGCTGTTCGTTCGTTCGGGGGTGAGGGGCAACTGTGTACCGTG 180
DB	1149 GTGACCGGTGTGTGCTGTGTGCTGTTCGTTCGTTCGGGGGTGAGGGGCAACTGTGTACCGTG 1208
OY	181 ATGTGATTCGGGCGCTACCGGGAACATGTGGAGAACACCACTTGTATCCTTGGGACGATG 240
DB	1209 ATGTGATTCGGGCGCTACCGGGAACATGTGGAGAACACCACTTGTATCCTTGGGACGATG 1268
OY	241 GCCGTGTCCGACTTACTGATCTTGTGTGGAGTGCAGTCGTTGACCTGTACCGCTCTGGCGC 300
DB	1269 GCCGTGTCCGACTTACTCATCTGTCTCGGGCTGCAGTCGTTGACCTGTACCGCTCTGGCGC 1328
OY	301 TTGGGGCCCTGGGTGTTGGGGCCGCTGCTGTGCGCCCTGTCCCTTACTGTGGGCGAAGGC 360
DB	1329 TTGGGGCCCTGGGTGTTGGGGCCGCTGCTGTGCGCCCTGTCCCTTACTGTGGGCGAAGGC 1388
OY	361 TGCACTTACGGCACGCTGTCTGCACATGACCCGCGCTACAGCTGTGAGCGCTTACCTGGCCATC 420
DB	1389 TGCACTTACGGCACGCTGTCTGCACATGACCCGCGCTACAGCTGTGAGCGCTTACCTGGCCATC 1448
OY	421 TGCCGCCGCGCTTCGGGCGCGCGTGTGTGTACAACCGGCGCGCGGTCCGCGCGCTATCGCT 480

Db	1449	TGCGCGCCGCTCCGCGCCGCGCTTGATGATACCCGCGCGCGCTCGCGCTCATGCGT	1508
Qy	481	GTGCTCTGGGCGGTGGCGCTGCTCTCGCGGTCCTCTTGTTCGTGTGGGCGTGGAG	540
Db	1509	GTGCTCTGGGCGGTGGCGCGTCTCTCGCGGTCCTCTTGTTCGTGTGGGCGTGGAG	1568
Qy	541	CAGGACCCCGGCAATCTCCGTAATCCCGGGGCTCAATGGCAACCGGGGGATACGCTCTCG	600
Db	1569	CAGGACCCCGGCAATCTCCGTAATCCCGGGGCTCAATGGCAACCGGGGGATACGCTCTCG	1628
Qy	601	CGTCTGCGCTCGTGGCGCGCTCTCTGGCTCTCGGGGCGGCAACCGCGTCCCGCGCTCG	660
Db	1629	CCTCTGCGCTCTGTCGCGCTCTCTTGCTCTCGGGGCGGCAACCGCGTCCCGCGCTCG	1688
Qy	661	GGGCGCGAAGCCGCGAAGCGCGCGCGCTGTTCAGCGCGAATGCGGCGAGCCCGCG	720
Db	1689	GGGCGCGAAGCCGCGAAGCGCGCGCGCTGTTCAGCGCGAATGCGGCGAGCCCGCG	1748
Qy	721	CAGCTGGGCGCGTGGCTGTCAATGTGTGGGTACCAACCGCTATCTTCTCTGCGCTTT	780
Db	1749	CAGCTGGGCGCGTGGCTGTCAATGTGTGGGTACCAACCGCTATCTTCTCTGCGCTTT	1808
Qy	781	CGTGTCTCAGCAATCCCTACGCGGCTCATGGGCGGGAGCTGTGAGACGCGCGCGCG	840
Db	1809	CTGTGCTCAGCAATCCCTACGCGGCTCATGGGCGGGAGCTGTGAGACGCGCGCGCG	1868
Qy	841	CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCCACCGCGAGACCAACCGCTCTCG	900
Db	1869	CTGCGAGCGCGCGCGCTCGGGCGGAGAGAGCCACCGCGAGACCAACCGCTCTCTG	1928
Qy	901	CGTAAGTAGAGCGCGGTGTTCAAAGAGCGCTGCTGCAATCGCGCGCGCGGAGCC	960
Db	1929	CGTAAGTAGAGCGCGGTGTTCAAAGAGCGCTGCTGCAATCGCGCGCGCGGAGCC	1988
Qy	961	GGCGAAACGCTGGGTGCCCTTCCCTGCTGCGCCAGCGTGTGGGGCGCGCTTCAAGTCCC	1020
Db	1989	GGCGAAACGCTGGGTGCCCTTCCCTGCTGCGCCAGCGTGTGGGGCGCGCTTCAAGTCCC	2048
Qy	1021	TTTCTATTTCGATTCAGAGCTCCAGCGCTCCAGCGCGGTACTTCCATCCCGCGAATAACATG	1080
Db	2049	--TCTATTTCGATTCAGAGCTCCAGCGCTCCAGCGCGGTACTTCCATCCCGCGAATAACATG	2106
Qy	1081	TCTGTGCCCGAGAGCTGTGGGGGAGCCCAAGGCGCTTTAGGGTGGATCCCCGATC	1140
Db	2107	TCTGTGCCCGAGAGCTGTGGGGGAGCCCAAGGCGCTTTAGGGTGGATCCCCGATC	2166
Qy	1141	CGATTAGTAAACAGAGTGTCTTTCAGAGCGCTGAGACCAATAAGGAAGATGTGATA	1200
Db	2167	CGATTAGTAAACAGAGTGTCTTTCAGAGCGCTGAGACCAATAAGGAAGATGTGATA	2226
Qy	1201	TTTCTAATCCAACACACTGTGTAGTGCACAAATAGAGAGTCTCACAGTCTCTTGA	1260
Db	2227	TTTCTAATCCAACACACTGTGTAGTGCACAAATAGAGAGTCTCTCACAGTCTCTTGA	2286
Qy	1261	AGAAGAGGAGATTCATTAAAGCTAAATTTTATTTAATGTTAAGTAGTGTGAAGGC	1320
Db	2287	AGAAGAGGAGATTCATTAAAGCTAAATTTTATTTAATGTTAAGTAGTGTGAAGGC	2346
Qy	1321	TAAAGTAAACCTTGCTGTGATCAAAAAGTAAAGATTGTGCAGACCTGTGTGAATCTT	1380
Db	2347	TAAAGTAAACCTTGCTGTGATCAAAAAGTAAAGATTGTGCAGACCTGTGTGAATCTT	2406
Qy	1381	TTTAAACAGAGAACAGAAAACCTGTCTCGAAGTGGGTTTGGAAGAGACCTGCAAG	1440
Db	2407	TTTAAACAGAGAACAGAAAACCTGTCTCGAAGTGGGTTTGGAAGAGACCTGCAAG	2466
Qy	1441	CGGCTGTTCAGAGAAATGTCTCTTCTGTGTATATGTTCAGCTTGATTAACAATATGG	1500
Db	2467	CGGCTGTTCAGAGAAATGTCTCTTCTGTGTATATGTTCAGCTTGATTAACAATATGG	2526
Qy	1501	AGCTTACTATGAGTTTAAAGCAAGTATCATGAGCTGAGCTGTGTCAATTTTCT	1566

Db	2527	AGCTACACATGACAGTTTAAAGCAAGTATCCATGACAGCTGACAGCTGATCTATTTTCT	2586
Qy	1561	GGGGTGAAGATGTGCGCTTGGTGAAGATTTTCTTAATTTATTTTCTGTACTTGTATTT	1620
Db	2587	GGGGTGAAGATGTGCGCTTGGTGAAGATTTTCTTAATTTATTTTCTGTACTTGTATTT	2646
Qy	1621	GCAGATGGTCTCTTGTCCGAGGGGAGGGAGTTATTTGCTTCCCAATGCTTTTGTATCCC	1680
Db	2647	GCAGATGGTCTCTTGTCCGAGGGGAGGGAGTTATTTGCTTCCCAATGCTTTTGTATCCC	2706
Qy	1681	GGTGTGTGTCTTATGTTGCAGTGGTGGTGTCTTCTGACATTTATAATTTGCTGTGGCCC	1740
Db	2707	GGTGTGTGTCTTATGTTGCAGTGGTGGTGTCTTCTGACATTTATAATTTGCTGTGGCCC	2766
Qy	1741	TTCCACGTTTGGCAGAAATCTATTACATAAACAAGAAAGATGGCGGATGATGACTTCTCT	1800
Db	2767	TTCCACGTTTGGCAGAAATCTATTACATAAACAAGAAAGATGGCGGATGATGACTTCTCT	2826
Qy	1801	CAGTACTTTAAATCATGTGCTGTGCAACTTTTCTATCTGAGGGCATCTATCAACCAATC	1860
Db	2827	CAGTACTTTAAATCATGTGCTGTGCAACTTTTCTATCTGAGGGCATCTATCAACCAATC	2886
Qy	1861	CTCTCAACCTTCATTTCAAGAAGTACAGAGCGGGGCTTTAACTGTGCTGCGAAGG	1920
Db	2887	CTCTCAACCTTCATTTCAAGAAGTACAGAGCGGGGCTTTAACTGTGCTGCGAAGG	2946
Qy	1921	AAGTCCAGGCGCAGAGGCTTCCACAGAACAGAGGACATCTGCGGGGAAATTCCAGGGGAC	1980
Db	2947	AAGTCCAGGCGCAGAGGCTTCCACAGAACAGAGGACATCTGCGGGGAAATTCCAGGGGAC	3006
Qy	1981	ACTGAGAGGAGACAGCGTGGGCTACACCGAGACAAACGCTAACGTGAAAGCATGGGATTA	2040
Db	3007	ACTGAGAGGAGACAGCGTGGGCTACACCGAGACAAACGCTAACGTGAAAGCATGGGATTA	3066
RESULT 3			
ID	AAF83684	AAF83684 standard; DNA; 1161 BP.	
XX	AAF83684;		
AC	23-JUL-2001	(first entry)	
XX			
DT			
XX			
DE	Short form of motilin receptor, GPR-38B isoform encoding DNA.		
XX			
KW	zsl333; signal transduction; hormone; enzyme; neural development;		
KW	gastric contractility; nutrient uptake; digestive; pancreatic; human;		
KW	insulin-like growth factor-I; growth hormone; bone; gastrointestinal;		
KW	glucose; osteopathic; anorectic; vulnerability; immunomodulator; GHS-R; ds;		
KW	G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..1161	
FT		/*tag= a	
FT		/product= "GPR-38B"	
XX			
PN	WO200138355-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	22-NOV-2000; 2000WO-US32074.		
XX			
PR	22-NOV-1999; 99US-0166765.		
XX			
PA	(ZYMO ) ZYMOGENETICS INC.		
XX			
PI	Sheppard PO, Jaepers SR, Deisher TA, Bishop PD;		
XX			
DR	WPI: 2001-355879/37.		
DR	P-PSDB; AAB62653.		
XX			











Query Match	43.9%;	Score 896.2;	DB 22;	Length 1239;
Best Local Similarity	99.7%;	Pred. No. 5.1e-174;		
Matches 898; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0

QY	1	ATGGGAGAGCCCTTGGAAACGGACAGACAGCCGCCCGAGAGGGGGGCGCGGAGACCGCCCGTGGCC	60
Db	1	ATGGGAGAGCCCTTGGAAACGGACAGACAGCCGCCCGAGAGGGGGGCGCGGAGACCGCCCGTGGCC	60
QY	61	GGCGTGGCGCCTTGGCGACGAGCGCGCGTGGCGCCCTTTCCCTGGGGGCGCTGGTGGC	120
Db	61	GGCGTGGCGCCTTGGCGACGAGCGCGCGTGGCGCCCTTTCCCTGGGGGCGCTGGTGGC	120
QY	121	GTGACCGCGTGTGGCTGTGGCTGTTCGTTCGTTCGGGGGTGAGCGGCAAGTGTGACCGTG	180
Db	121	GTGACCGCGTGTGTGGCTGTGGCTGTTCGTTCGTTCGGGGGTGAGCGGCAAGTGTGACCGTG	180
QY	181	ATGCTGATGGGCGGCTACCGGGACATGGCGAACACACACAACTTGTATCTGGGAGCATG	240
Db	181	ATGCTGATGGGCGGCTACCGGGACATGGCGAACACACACAACTTGTATCTGGGAGCATG	240
QY	241	GGCGGTGCGACACTTACTACTCTGCTCGGAGCTGGCGTTGACCTGTACCGCCTTGCGGC	300
Db	241	GGCGGTGCGACACTTACTACTCTGCTCGGAGCTGGCGTTGACCTGTGTACCGCCTTGCGGC	300
QY	301	TGGCGAGCCCTGGGATTTGGGGCGCGCTGGCTGTGCTGTGCGCGCTGTCCCTCTAAGTGGAGAGG	360
Db	301	TGGCGAGCCCTGGGATTTGGGGCGCGCTGGCTGTGCTGTGCGCGCTGTCCCTCTAAGTGGAGAGG	360
QY	361	TGGACCTTACGCCACGCTGTGTGACATGACCGCGCTCAGCGTGCAGCGCTAACCTGGGCATC	420
Db	361	TGGACCTTACGCCACGCTGTGTGACATGACCGCGCTCAGCGTGCAGCGCTAACCTGGGCATC	420
QY	421	TGCGCCCGCGCTCCGAGCGCGCTCTTGGTCAACCGGCGCGCGTGCACCGCGCTACCTGGGCATC	480
Db	421	TGCGCCCGCGCTCCGAGCGCGCTCTTGGTCAACCGGCGCGCGTGCACCGCGCTACCTGGGCATC	480
QY	481	GTGCTTGTGGAGCGTGGCGCTGCTCTGTGCGGATCCCTTTCTTGTCTGTGTGGCGTGGAG	540
Db	481	GTGCTTGTGGAGCGTGGCGCTGCTCTGTGCGGATCCCTTTCTTGTCTGTGTGGCGTGGAG	540
QY	541	CAGGACCCCGGATGTCGGTATGATCCCGGGGCGCTCAATGGCAACGGCGGAGATGCGCTCTCG	600
Db	541	CAGGACCCCGGATGTCGGTATGATCCCGGGGCGCTCAATGGCAACGGCGGAGATGCGCTCTCG	600
QY	601	CCTCTGCGCTGTGCGCGCGCTCTGTGAGCTCTGCGGGGCGCACCGCGTCCCGCGCGTGG	660
Db	601	CCTCTGCGCTGTGCGCGCGCTCTGTGAGCTCTGCGGGGCGCACCGCGTCCCGCGCGTGG	660
QY	661	GGGCGCCGAGATCCGCGGAGAGCGCGCGCGCTGTTCAGCCGCAATGCGCGCGAGCGCTCGCG	720
Db	661	GGGCGCCGAGATCCGCGGAGAGCGCGCGCGCTGTTCAGCCGCGCAATGCGCGCGAGCGCTCGCG	720
QY	721	CAGCTGAGGCGGCTGCGTGTGATGCTGTGGGTCAACACCGCGCTACTTCTTCGCGCTTT	780
Db	721	CAGCTGAGGCGGCTGCGTGTGATGCTGTGGGTCAACACCGCGCTACTTCTTCGCGCTTT	780
QY	781	CTGTGCTTCAAGATCTCTTACGGGCTCATCGGGCGGAGCTGTGTGAGCAAGCGCGCGGCGG	840
Db	781	CTGTGCTTCAAGATCTCTTACGGGCTCATCGGGCGGAGCTGTGTGAGCAACCGCGCGGCGG	840
QY	841	CTGCGAGGCGCGGCGCGCTCTCGGGGCGGAGAGAGAGCCACCGGCGAGCAACAAAGCGTCTGG	900
Db	841	CTGCGAGGCGCGGCGCGCTCTCGGGGCGGAGAGAGAGCCACCGGCGAGCAACCGCGGCTCTGG	900
QY	901	C 901	
Db	901	C 901	

XX	AAI66989;
AC	
XX	30-JAN-2002 (first entry)
XX	
DT	
XX	Human GPR38 variant GPR38V polypeptide encoding DNA.
XX	
KW	GPR38V; variant; antibacterial; cytosstatic; analgesic; antiasthmatic;
KM	anti-Parkinsonian; hypertensive; hypotensive; antidiabetic; osteopathic;
KX	antiallergic; antimigraine; neuroleptic; nootropic; anticonvulsant;
KW	ant ulcer; antiemetic; cardiant; vaccine; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..1506
FT	/tag= a
FT	/product= "GPR38V"
PN	WO200164836-A2.
PD	
XX	07-SEP-2001.
PF	
XX	28-FEB-2001; 2001WO-US06277.
PR	
XX	01-MAR-2000; 2000US-0516315.
PA	(SMIK ) SMITHKLINE BEECHAM CORP.
PI	
XX	Elehoubagy N, Shabon U;
DR	
XX	WPI: 2001-638956/73.
P	P-PSDB; AAG65822.
XX	
PT	New human GPR38V polypeptide and polymucleotide, useful for treating
PT	e.g. bacterial, fungal, protozoal and viral infections, cancers or
PT	allergies, as vaccines, and for identifying agonists and antagonists
PT	potentially useful in therapy -
XX	
PS	Claim 2; Page 26; 32pp; English.
XX	
CC	This DNA encodes a human GPR38 variant (GPR38V) polypeptide. GPR38V can
CC	be expressed by standard recombinant methodology. The polymucleotides and
CC	polypeptides are used in the treatment of bacterial, fungal, protozoal
CC	and viral infections, infections caused by HIV-1 or HIV-2, pain, cancers,
CC	diseases, obesity, anorexia, asthma, Parkinson's disease, acute heart
CC	failure, hypertension, urinary retentions, osteoporosis, allergies,
CC	ulcers, migraine, psychotic and neurological disorders, or dyskinesias.
CC	They are also useful for identifying agonists and antagonists that are
CC	potentially useful in therapy, as vaccines to induce immunological
CC	response in a mammal. The polypeptides may also be used as immunogens to
CC	produce antibodies immunospecific for the polypeptides, and to identify
CC	membrane bound or soluble receptors.
XX	
SQ	Sequence 1506 BP; 207 A; 534 C; 509 G; 256 T; 0 other;

Query March	43.9%	Score 896.2;	DB 22;	Length 1506;
Best Local Similarity	99.7%;	Pred. No. 5.3e-174;		
Matches 898;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

Dy	1	ATGGGACAGCCCCCTTGGAACGCGACGCCAGCCCAGGGGGCGCGGAGCCGCGTGGCCCC	60
Db	268	ATGGGCAACCCCTTGGAACGCGACGCGAGCGGCCCCGAGGGGGCGCGGAGCCGCGTGGCCCC	327
Dy	61	GCGGTGCGGCCTTGGAAGAGAGCGCGCGTGTGTCCTTTCCCTGCGGGCGCTGTGTCCG	120
Db	328	GCGGTGCGGCCTTGGAAGAGAGCGCGCGTGTGTCCTTTCCCTGCGGGCGCTGTGTCCG	387
Dy	121	GTGACCGCTGTGTGCGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	180
Db	388	GTGACCGCTGTGTGCTGT	447
Dy	181	ATGCTGATGGGCGCGTACGGGAGCATGCGGACCAACCAACTTGTACTGGGCGAGCATG	240

Db 448 ATGCGATCGGCGCTACCGGAGATGCGGACACACCAACTTGTACTGGGACAGATG 507  
 Qy 241 GCCGTGTCGAGACTTACTACTCTGCTCGGAGCTGCCCTTGAACCTGACCGCTTGGCG 300  
 Db 508 GCCGTGTCGAGACTTACTACTCTGCTCGGAGCTGCCCTTGAACCTGACCGCTTGGCG 567  
 Qy 301 TCGGCGCCCTGGGATGTTGGGCGCGCTGCTCGGCGCTGCTGCTTACTGAGTGGGAGG 360  
 Db 568 TCGGCGCCCTGGGATGTTGGGCGCGCTGCTCGGCGCTGCTGCTTACTGAGTGGGAGG 627  
 Qy 361 TGCACCTACGCCAGCTGCTGTGACATGACCGCGCTCAGCGTGAGCGCTTACTGGCCATC 420  
 Db 628 TGCACCTACGCCAGCTGCTGTGACATGACCGCGCTCAGCGTGAGCGCTTACTGGCCATC 687  
 Qy 421 TCGCGCGCGCTCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Db 688 TCGCGCGCGCTCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747  
 Qy 481 GTGCTGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 Db 748 GTGCTGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807  
 Qy 541 CAGGACCCCGGAGATCTCCGTAGTCCCGGCGCTTCAATGACACCGCGCGGATGCTCTCG 600  
 Db 808 CAGGACCCCGGAGATCTCCGTAGTCCCGGCGCTTCAATGACACCGCGCGGATGCTCTCG 867  
 Qy 601 CCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 Db 868 CCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927  
 Qy 661 GGGGCGGAGCGCGGAGCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 Db 928 GGGGCGGAGCGCGGAGCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987  
 Qy 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 Db 988 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047  
 Qy 781 CTGTGCTTACGACTCTCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTT 840  
 Db 1048 CTGTGCTTACGACTCTCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTT 1107  
 Qy 841 CTGCGAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 Db 1108 CTGCGAGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167  
 Qy 901 C 901  
 Db 1168 C 1168  
 RESULT 9  
 AAF85448  
 ID AAF85448 standard; cDNA; 1203 BP.  
 XX  
 AC AAF85448;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE Nucleotide sequence of a rabbit motilin receptor polypeptide.  
 XX  
 KW Motilin receptor; gastrointestinal disease; gastric motility disorder;  
 XX gastroparesis; irritable bowel syndrome; diarrhoea; ss.  
 OS Oryctolagus cuniculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1203  
 FT /tag= a  
 FT /note= "motilin receptor"  
 XX  
 PN WO200132710-A1.

XX 10-MAY-2001.  
 PD 25-OCT-2000; 2000WO-US29426.  
 XX  
 PF 29-OCT-1999; 99US-0162264.  
 PR  
 XX (MERI) MERCK & CO INC.  
 PA  
 PI Tan C, McKee K;  
 XX  
 DR WPI: 2001-343479/36.  
 DR P-PSDB; AAB68477.  
 PT Novel polypeptides related to dog and rabbit motilin receptor  
 PT polypeptide, comprising unique regions from dog and motilin receptor  
 PT amino acid sequence, useful for identifying compounds for treating  
 PT diarrhoea in humans  
 XX  
 PS Claim 18; Page 18-19; 42bp; English.  
 CC The present sequence encodes a rabbit motilin receptor polypeptide.  
 CC The specification describes a unique sequence present in exon 1 of  
 CC the dog motilin receptor, which is not present in human or Sphaeroides  
 CC nephelus 7557 motilin receptor sequences. The unique nucleic acid  
 CC sequence is useful for measuring the ability of a compound to affect  
 CC motilin receptor activity. Motilin receptor polynucleotides and  
 CC polypeptides are used to identify therapeutic compounds which are  
 CC useful for treating gastrointestinal diseases and disorders such as  
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,  
 CC and diarrhoea.  
 XX  
 SQ Sequence 1203 BP; 154 A; 423 G; 223 T; 0 other;  
 XX  
 Query Match 32.8%; Score 669.8; DB 22; Length 1203;  
 Best Local Similarity 86.3%; Pred. No. 1.1e-127;  
 Matches 778; Conservative 0; Mismatches 87; Indels 36; Gaps 2;  
 Qy 1 ATGGGACAGCCCTTGGAGACGAGCGACCGCCCGAGGAGGCGCGGAGCGCGCGCGCGCGCG 60  
 Db 1 ATGGGACAGCCCTTGGAGACGAGCGACGCGCCCGAGGAGCGCGCGGAGCGCGCGCGCGCGCG 60  
 Qy 61 GCGCTGCGCGCTTGGAGAGCGCGCGCTGCGCGCTTCCCTGAGGAGCGCTGAGCGCG 120  
 Db 61 GCGCTGCGCGCTTGGAGAGCGCGCGCTGCGCGCTTCCCTGAGGAGCGCTGAGCGCG 120  
 Qy 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 Db 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 Qy 181 ATGCTGATCGGCGCTGACCGGAGCATGCGGACACACCACTTGTACTGGGACAGATG 240  
 Db 181 CTGCTGATCGGCGCTGACCGGAGCATGCGGACACACCACTTGTACTGGGACAGATG 240  
 Qy 241 GCCGTGTCGAGACTTACTACTCTGCTCGGAGCTGCCCTTGAACCTGACCGCTTGGCG 300  
 Db 241 GCCGTGTCGAGACTTACTACTCTGCTCGGAGCTGCCCTTGAACCTGACCGCTTGGCG 300  
 Qy 301 TCGGCGCCCTGGGATGTTGGGCGCGCTGCTCGGCGCTGCTGCTTACTGAGTGGGAGG 360  
 Db 301 TCGGCGCCCTGGGATGTTGGGCGCGCTGCTCGGCGCTGCTGCTTACTGAGTGGGAGG 360  
 Qy 361 TGCACCTACGCCAGCTGCTGTGACATGACCGCGCTCAGCGTGAGCGCTTACTGGCCATC 420  
 Db 361 TGCACCTACGCCAGCTGCTGTGACATGACCGCGCTCAGCGTGAGCGCTTACTGGCCATC 420  
 Qy 421 TCGCGCGCGCTCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Db 421 TCGCGCGCGCTCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Qy 481 GTGCTGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 Db 481 GCGCTGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 541 CAGACCCCGCATCTCCGTAATCCCGGACCTTCATAGCACCGCGCGATCGCTCTCTCG 600  
 Db 541 CAGACCCCGCGCTCTTCGCGGACCCCGGACCGCAACGTAATGTGCGCTGACCCCTCG 600  
 QY 601 CCTTCGCGCTGTGCGCGGCTCTTCGCGCTCTCGCGGCGCCACCGCGCTCTCGCGCTCG 660  
 Db 601 TCGCCCGCC-----CGGCGTCTCCCGCGCTCG 627  
 QY 661 GGGCCCGACCGCGAGCGCGCGCGCTTCAGCCCGCAATGCGCGCGCGCGCGCG 720  
 Db 628 GGGCCCG---GAGCCGAGGCGCGCGCTCTTCAGCCCGAGTCCGCGCGCGCGCGCG 664  
 QY 721 CAGCTGGGCGCGCTGCGGTCTCATGCTGTGGGTACCAACCGCTCTCTTCCTGCGCGCTTT 780  
 Db 685 CAGCTGGGCGCTGCGGTCTCATGCTGTGGGTACCAACCGCTCTCTTCCTGCGCGCTTC 744  
 QY 781 CTGCGCTCAGCATCTCTTCAGCGGCTCATCGCGCGGAGCTGTGGAGCAGCGCGCGCG 840  
 Db 745 CTGCGCTCAGCATCTCTTCAGCGGCTCATCGCGCGGAGCTGTGGAGGCTGTGGAGCGCG 804  
 QY 841 CTGCGAGCGCGCGCGCGCTTCGCGGCGGAGAGAGCGCAACCGCGCAACCGCGCTTCG 900  
 Db 805 CTGCGAGCGCGCGCGCGCGCGCGCTTCGCGGAGAGAGCGCGCAACCGCGCAACCGCGCTTCG 864  
 QY 901 C 901  
 Db 865 C 865

## RESULT 10

AAF85447  
 ID AAF85447 standard; DNA; 813 BP.

AAF85447;

23-JUL-2001 (first entry)

Nucleotide sequence of a dog motilin receptor exon 1.

Motilin receptor; gastrointestinal disease; gastric motility disorder;  
 gastroparesis; irritable bowel syndrome; diarrhoea, ss.

Canis sp.

Location/Qualifiers

CDS 1..813  
 /+tag= a  
 /note= "motilin receptor exon 1"

W0200132710-A1.

10-MAY-2001.

25-OCT-2000; 2000MO-US29426.

29-OCT-1999; 99US-0162264.

(MERI) MERCK & CO INC.

Tan C, McKee K;

WPI, 2001-343479/36.

P-PSDB; AAB68476.

Novel polypeptides related to dog and rabbit motilin receptor  
 polypeptide, comprising unique regions from dog and motilin receptor  
 amino acid sequence, useful for identifying compounds for treating  
 diarrhoea in humans

Claim 14; Page 17-18; 42pp; English.

The present sequence represents exon 1 of a dog motilin receptor gene.

CC The specification describes an unique sequence present in exon 1 of  
 CC the motilin receptor, which is not present in human or Spharoides  
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid  
 CC sequence is useful for measuring the ability of a compound to affect  
 CC motilin receptor activity. Motilin receptor polynucleotides and  
 CC polypeptides are used to identify therapeutic compounds which are  
 CC useful for treating gastrointestinal diseases and disorders such as  
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,  
 CC and diarrhoea.  
 XX

Sequence 813 BP; 65 A; 316 C; 305 G; 127 T; 0 other;

Query Match 26.5%; Score 541.2; DB 22; Length 813;

Best Local Similarity 80.1%; Pred. No. 2.1e-101;

Matches 702; Conservative 0; Mismatches 108; Indels 66; Gaps 3;

QY 25 GAGGAGCCCGAGGAGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 84  
 Db 4 GAGGAGCCCGAGGAGGAGCGCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCG 63  
 QY 85 CGCTGCTCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 144  
 Db 64 CTGCTGCTCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 123  
 QY 145 TTGCTGCTCGGAGTGAAGCGGCAACGTGTGACCTGTGATGAGGCGCTACCGGAGC 204  
 Db 124 TTGCGCGTGGCGGAGCGGCAACGTGTGACCTGTGATGAGGCGCTACCGGAGC 183  
 QY 205 ATGCGAGCACACCAACCTTGTACCTTGGGAGCATGAGCGCTGTGCGACTTACATCTTG 264  
 Db 184 ATGCGAGCACACCAACCTTGTACCTTGGGAGCATGAGCGCTGTGCGACTTACATCTTG 243  
 QY 265 CTGCGGCGCTGCTTCGACCTGTACCGCGCTTCGCGGCGCTTCGCGGCGCTTCGCGGCG 324  
 Db 244 CTGCGGCGCTGCTTCGACCTGTACCGCGCTTCGCGGCGCTTCGCGGCGCTTCGCGGCG 303  
 QY 325 CTGCTGCTCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 384  
 Db 304 CTGCTGCTCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 363  
 QY 385 ATGACCGCGCTACGCGTTCGAGCGCGTACCTTGGCAATCTGCGCGCGCGCTTCGCGCG 444  
 Db 364 GTGACCGCGCTGAGCGCGTACCTTGGCAATCTGCGCGCGCGCTTCGCGCGCGCGCG 423  
 QY 445 TTGCTACCG 504  
 Db 424 CTGCTGCTCG 483  
 QY 505 TCTGCGGATCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 564  
 Db 484 TCTGCGGCGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 534  
 QY 565 CCGGAGCTCAATGAGCACCGGCGGAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 624  
 Db 535 CCGGAGCTCAACGAGCGCGCG----- 557  
 QY 625 TGGCTCTCGCGGCG 684  
 Db 558 -----GCTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597  
 QY 685 GCGCTGTTAGCGCGCAATGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 744  
 Db 598 GCGCTGTTAGCGCGAGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657  
 QY 745 CTGCTGTTAGCG 804  
 Db 658 CTGCTGTTAGCG 717  
 QY 805 CTCATCGCGCGGAGCTGTGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 864  
 Db 718 CGCATTCGCGCGGAGCTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777  
 QY 865 CGGAGAGAGCGCACCGCGGAGACCAACCGCGCTCTG 900















GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:58 ; Search time 5959 Seconds

(without alignments)  
9963.035 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040  
Sequence: 1 atgggagagccctcgaacgg.....acgtgaagacgatggatata 2040

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_da.\*  
2: gb\_hlg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pac.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_hlg\_hum.\*  
31: em\_hlg\_inv.\*  
32: em\_hlg\_other.\*  
33: em\_hlg\_mus.\*  
34: em\_hlg\_pln.\*  
35: em\_hlg\_rod.\*  
36: em\_hlg\_man.\*  
37: em\_hlg\_vtc.\*  
38: em\_sy.\*  
39: em\_hlg\_hum.\*  
40: em\_hlg\_mus.\*  
41: em\_hlg\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2035.2	99.8	2040	9	AF034632	AF034632 Homo sapi
2	2033.6	99.7	163284	9	AL137000	AL137000 Human DNA
3	2033.6	99.7	341560	2	AL596304	AL596304 Homo sapi
4	1047.2	51.3	1161	6	AX154591	AX154591 Sequence
5	896.2	43.9	1239	6	AX154589	AX154589 Sequence
6	592.2	29.0	692	9	HS339407	AJ339407 Homo sapi
7	572.2	28.0	692	9	HS342408	AJ342408 Homo sapi
8	557.4	27.3	615	9	HS339459	AJ339459 Homo sapi
9	424.8	20.8	701	9	HS3326768	AJ3326768 Homo sapi
10	307	15.0	172168	2	AC109158	AC109158 Mus muscu
11	283	13.9	283	6	AR186467	AR186467 Sequence
12	283	13.9	283	6	AR182284	AR182284 Sequence
13	266.8	13.1	573	9	HS341295	AJ341295 Homo sapi
14	234.8	11.5	1063	6	AR156351	AR156351 Sequence
15	234.8	11.5	1101	4	SSU60178	U60178 Sus scrofa
16	233.2	11.4	870	4	SSU60180	U60180 Sus scrofa
17	230	11.3	1029	6	AR156352	AR156352 Sequence
18	230	11.3	142736	2	AC117349	AC117349 Rattus no
19	230	11.3	169279	2	AC095173	AC095173 Rattus no
20	230	11.3	181578	2	AC130771	AC130771 Rattus no
21	230	11.3	195125	2	AC108400	AC108400 Mus muscu
22	230	11.3	227035	2	AC116585	AC116585 Mus muscu
23	228.8	11.2	870	9	HSU60181	U60181 Human growt
24	228.8	11.2	1122	6	AR156354	AR156354 Sequence
25	228.8	11.2	6787	9	AF369786	AF369786 Homo sapi
26	228.8	11.2	145593	9	AC069523	AC069523 Homo sapi
27	228.4	11.2	1092	6	AR156357	AR156357 Sequence
28	228.4	11.2	1350	10	AB001382	AB001382 Rattus no
29	228.4	11.2	3129	6	AR156356	AR156356 Sequence
30	228.4	11.2	3129	10	RMU94321	U94321 Rattus norv
31	227.8	11.2	1088	6	AR156353	AR156353 Sequence
32	227.8	11.2	1101	6	AX154584	AX154584 Sequence
33	227.8	11.2	1101	9	HSU60179	U60179 Human growt
34	225.2	11.0	1676	5	AF082210	AF082210 Spinoeroid
35	221	10.8	560	10	AY056474	AY056474 Mus muscu
36	216	10.6	1233	4	AY093950	AY093950 Ovis arie
37	216	10.6	1591	4	AY093949	AY093949 Ovis arie
38	216	10.6	1816	4	AY093948	AY093948 Ovis arie
39	214	10.5	1455	5	AF082209	AF082209 Spinoeroid
40	194.4	9.5	557	10	AF332997	AF332997 Mus muscu
41	184.6	9.0	1098	6	AX154586	AX154586 Sequence
42	176.8	8.7	528	11	G68328	G68328 GHR X2.3 H
43	161.4	7.9	130311	2	AC095360	AC095360 Rattus no
44	160.4	7.9	130311	2	AC095360	AC095360 Rattus no
45	153	7.5	836	6	AR156355	AR156355 Sequence

## ALIGNMENTS

RESULT 1  
AF034632 2040 bp DNA linear PRI 13-JUL-1998  
LOCUS Homo sapiens orphan G protein-coupled receptor (GPR38) gene,  
DEFINITION complete cds.  
ACCESSION AF034632  
VERSION AF034632.1 GI:2654158  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2040)  
McKee,K.K., Tan,C.P., Palya,O.C., Liu,U., Feigmer,S.D.,  
Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,L.H.

TITLE Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors

JOURNAL J. Biol. Chem. 278:9417-9426 (1997)

MEDLINE 941746

PUBMED 941746

REFERENCES 2 (bases 1 to 2040)

AUTHORS McKee, K.K., Tan, C.P., Palyha, O.C., Liu, J., Feighner, S.D., Hreniuk, D.L., Smith, R.G., Van Der Ploeg, L.H.T. and Howard, A.D.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co., Inc., PO Box 2000, Rahway, NJ 07065, USA

FEATURES

source 1..2040

organism="Homo sapiens"

db\_xref="taxon:9606"

chromosome="13"

map="13q14-21"

<1..>2040

gene /gene="GPR38"

mRNA join(<1..901,1703..>2040)

CDS join(1..901,1703..2040)

/gene="GPR38"

/note="orphan G protein-coupled receptor related to growth hormone secretagogue and neurotensin receptors"

/codon\_start=1

/product="GPR38"

/protein\_id="AAC26081.1"

/db\_xref="GI:2654159"

/translation="MGSPWNGSDPEGRPPPALPPCDERRCSPEPLGALVPTAV  
CLCLFVGVGNAVVTVMILIGRYDMRTTTLVYGSMAVSPLLILGLPDLRLMR  
PVMYGLICRLSTVYGGCTATLHMTALSVRYALICPLARVLTBRVRLALIA  
VLMAYALLSAGPLFLVGYRQDPSISVPLNGSTARSPLASPLMSPPSP  
PSGPETAALALFSLRECRSPADJGALVLMVMTATVFPFLCLSLTYLIGRELS  
SRPLRGPASGRGRQTVRLVLAFLICWLPFHVRILTYINTDSRMVTFEQ  
YENVALQFLYLSASINPILYNLISKVYRAAPFLILARKSRPRGFRSDTAGEVAG  
DTGGTVGTETSAVNTMG"

BASE COUNT 347 a 631 c 598 g 464 t

ORIGIN

Query Match 99.8%; Score 2035.2; DB 9; Length 2040;

Best Local Similarity 99.3%; Pred. No. 1e-304;

Matches 2037; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCGAGCCCTTGGAACGCGAGCGAGCGCCCGAGGGGGGGGGGGAGCCCGCTGGGCC 60

DB 1 ATGGGCGAGCCCTTGGAACGCGAGCGAGCGCCCGAGGGGGGGGGGGAGCCCGCTGGGCC 60

QY 61 GCGCTGCGGCTTGCGAGCGAGCGCGCTGCGCCCTTCCCTGGGGGCGCTGGTGGCG 120

DB 61 GCGCTGCGGCTTGCGAGCGAGCGCGCTGCGCCCTTCCCTGGGGGCGCTGGTGGCG 120

QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCGGGGTGAGCGGCAACGTGTGACCGGTG 180

DB 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCGGGGTGAGCGGCAACGTGTGACCGGTG 180

QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCGGGGTGAGCGGCAACGTGTGACCGGTG 180

DB 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCGGGGTGAGCGGCAACGTGTGACCGGTG 180

QY 181 ATGTGATCGGGGCGCTACCGGGACATGCGACCAACCACTTGTAAGTGGGCGCATG 240

DB 181 ATGTGATCGGGGCGCTACCGGGACATGCGACCAACCACTTGTAAGTGGGCGCATG 240

QY 181 ATGTGATCGGGGCGCTACCGGGACATGCGACCAACCACTTGTAAGTGGGCGCATG 240

DB 181 ATGTGATCGGGGCGCTACCGGGACATGCGACCAACCACTTGTAAGTGGGCGCATG 240

QY 241 GCGGTGCGGACCTACTCATCTGTGCTGGGGTGGCGCTGCACTGTACCGCTTGGGGCG 300

DB 241 GCGGTGCGGACCTACTCATCTGTGCTGGGGTGGCGCTGCACTGTACCGCTTGGGGCG 300

QY 241 GCGGTGCGGACCTACTCATCTGTGCTGGGGTGGCGCTGCACTGTACCGCTTGGGGCG 300

DB 241 GCGGTGCGGACCTACTCATCTGTGCTGGGGTGGCGCTGCACTGTACCGCTTGGGGCG 300

QY 301 TCGGGGCGCTGGGTGTTCGGGGCGGCTGTGCGCGCTGTCTTACGTGGGCGAGGCG 360

DB 301 TCGGGGCGCTGGGTGTTCGGGGCGGCTGTGCGCGCTGTCTTACGTGGGCGAGGCG 360

QY 301 TCGGGGCGCTGGGTGTTCGGGGCGGCTGTGCGCGCTGTCTTACGTGGGCGAGGCG 360

DB 301 TCGGGGCGCTGGGTGTTCGGGGCGGCTGTGCGCGCTGTCTTACGTGGGCGAGGCG 360

QY 361 TGCACCTACGCGCAGCTGTGCAATGACCGGCTCAGCGCTGACCGCTTGGCCATC 420

DB 361 TGCACCTACGCGCAGCTGTGCAATGACCGGCTCAGCGCTGACCGCTTGGCCATC 420

QY 421 TGGCCCGCGCTCGCGCGCGCGCTGTGTGTCACCGGGCGCGCGCTCGCGCGCTCATCGCT 480

DB 421 TGGCCCGCGCTCGCGCGCGCGCGCTGTGTGTCACCGGGCGCGCGCTCGCGCGCTCATCGCT 480

QY 481 GTGCTGTGGGCGCGCGCGCTGTGTGCTGTGCGCGCGCGCTGTGTGCTGTGTGCGCGCG 540

DB 481 GTGCTGTGGGCGCGCGCGCTGTGTGCTGTGCGCGCGCGCTGTGTGCTGTGTGCGCGCG 540

QY 541 CAGAGCCCGCGCATCTCGTAGTCCCGGAGCTCAATGAGCAGCGCGAGATCGCTCTCG 600

DB 541 CAGAGCCCGCGCATCTCGTAGTCCCGGAGCTCAATGAGCAGCGCGAGATCGCTCTCG 600

QY 541 CAGAGCCCGCGCATCTCGTAGTCCCGGAGCTCAATGAGCAGCGCGAGATCGCTCTCG 600

DB 541 CAGAGCCCGCGCATCTCGTAGTCCCGGAGCTCAATGAGCAGCGCGAGATCGCTCTCG 600

QY 601 CCTCTCGCGCTGTGCGCGCGCTCTCTGTGCTGTGCGCGCGCGCAACCGCTTCCCGCG 660

DB 601 CCTCTCGCGCTGTGCGCGCGCTCTCTGTGCTGTGCGCGCGCGCAACCGCTTCCCGCG 660

QY 601 CCTCTCGCGCTGTGCGCGCGCTCTCTGTGCTGTGCGCGCGCGCAACCGCTTCCCGCG 660

DB 601 CCTCTCGCGCTGTGCGCGCGCTCTCTGTGCTGTGCGCGCGCGCAACCGCTTCCCGCG 660

QY 661 GGGCGCGAGCGCGGAGGCGCGCGCGCGCTGTGCGCGCGGAAATGCGCGCGAGCGCGCG 720

DB 661 GGGCGCGAGCGCGGAGGCGCGCGCGCGCTGTGCGCGCGGAAATGCGCGCGAGCGCGCG 720

QY 661 GGGCGCGAGCGCGGAGGCGCGCGCGCGCTGTGCGCGCGGAAATGCGCGCGAGCGCGCG 720

DB 661 GGGCGCGAGCGCGGAGGCGCGCGCGCGCTGTGCGCGCGGAAATGCGCGCGAGCGCGCG 720

QY 721 CAGCTGGGCGCGCTGTGTCATGCTGTGAGTCAACACCGCGCTACTTCTTCCGCTTT 780

DB 721 CAGCTGGGCGCGCTGTGTCATGCTGTGAGTCAACACCGCGCTACTTCTTCCGCTTT 780

QY 721 CAGCTGGGCGCGCTGTGTCATGCTGTGAGTCAACACCGCGCTACTTCTTCCGCTTT 780

DB 721 CAGCTGGGCGCGCTGTGTCATGCTGTGAGTCAACACCGCGCTACTTCTTCCGCTTT 780

QY 781 CTGTGCTTCAAGCATCTCTTACGAGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCG 840

DB 781 CTGTGCTTCAAGCATCTCTTACGAGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCG 840

QY 781 CTGTGCTTCAAGCATCTCTTACGAGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCG 840

DB 781 CTGTGCTTCAAGCATCTCTTACGAGGCTCATCGGGCGGAGCTGTGAGCAGCGCGCGCG 840

QY 841 CTGCGAGCGCGCGCGCGCGCGCGCGCGGAGAGAGGCAACCGGAGACGCGCTCTG 900

DB 841 CTGCGAGCGCGCGCGCGCGCGCGCGCGGAGAGAGGCAACCGGAGACGCGCTCTG 900

QY 841 CTGCGAGCGCGCGCGCGCGCGCGCGCGGAGAGAGGCAACCGGAGACGCGCTCTG 900

DB 841 CTGCGAGCGCGCGCGCGCGCGCGCGCGGAGAGAGGCAACCGGAGACGCGCTCTG 900

QY 901 CGTAGGTGAGCGCGCGCGCTGTGTCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 960

DB 901 CGTAGGTGAGCGCGCGCGCTGTGTCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 960

QY 901 CGTAGGTGAGCGCGCGCGCTGTGTCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 960

DB 901 CGTAGGTGAGCGCGCGCGCTGTGTCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 960

QY 961 GCGCAAGCGCTGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

DB 961 GCGCAAGCGCTGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 961 GCGCAAGCGCTGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

DB 961 GCGCAAGCGCTGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 1021 TTTTCTATTGATTCAGAGCTTCCAGCGCGCGGTAAGCTTCCATCCCGGAGAAACATG 1080

DB 1021 TTTTCTATTGATTCAGAGCTTCCAGCGCGCGGTAAGCTTCCATCCCGGAGAAACATG 1080

QY 1021 TTTTCTATTGATTCAGAGCTTCCAGCGCGCGGTAAGCTTCCATCCCGGAGAAACATG 1080

DB 1021 TTTTCTATTGATTCAGAGCTTCCAGCGCGCGGTAAGCTTCCATCCCGGAGAAACATG 1080

QY 1081 TCTCTTCCCGCAGAGCTCTGTGGGGAGCCCAAGGCGCTTGAAGGTGGGATCCCGGATC 1140

DB 1081 TCTCTTCCCGCAGAGCTCTGTGGGGAGCCCAAGGCGCTTGAAGGTGGGATCCCGGATC 1140

QY 1081 TCTCTTCCCGCAGAGCTCTGTGGGGAGCCCAAGGCGCTTGAAGGTGGGATCCCGGATC 1140

DB 1081 TCTCTTCCCGCAGAGCTCTGTGGGGAGCCCAAGGCGCTTGAAGGTGGGATCCCGGATC 1140

QY 1141 CGATTCAAGTACCAAGTGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 1200

DB 1141 CGATTCAAGTACCAAGTGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 1200

QY 1141 CGATTCAAGTACCAAGTGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 1200

DB 1141 CGATTCAAGTACCAAGTGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 1200

QY 1201 TTTCTTAATCCACCAACCTGTGATGSCCAAAATGAGAGCTCTCAGAGCTCTTGAGA 1260

DB 1201 TTTCTTAATCCACCAACCTGTGATGSCCAAAATGAGAGCTCTCAGAGCTCTTGAGA 1260

QY 1201 TTTCTTAATCCACCAACCTGTGATGSCCAAAATGAGAGCTCTCAGAGCTCTTGAGA 1260

DB 1201 TTTCTTAATCCACCAACCTGTGATGSCCAAAATGAGAGCTCTCAGAGCTCTTGAGA 1260

QY 1261 AGACAGGAGATTTCAATTAAGCTAAATTTTATTAATGATTAATGATGCTGAAGGC 1320

DB 1261 AGACAGGAGATTTCAATTAAGCTAAATTTTATTAATGATTAATGATGCTGAAGGC 1320

QY 1261 AGACAGGAGATTTCAATTAAGCTAAATTTTATTAATGATTAATGATGCTGAAGGC 1320

DB 1261 AGACAGGAGATTTCAATTAAGCTAAATTTTATTAATGATTAATGATGCTGAAGGC 1320

QY 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGTGACAGCTGTGTAAATTTCT 1380

DB 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGTGACAGCTGTGTAAATTTCT 1380

QY 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGTGACAGCTGTGTAAATTTCT 1380

DB 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTGTGACAGCTGTGTAAATTTCT 1380

QY 1381 TTTCAACAGAGAACAGAAACCTTGTCTGAGAGTGGTTGTGAGAGAGAGCTGCCAAG 1440

DB 1381 TTTCAACAGAGAACAGAAACCTTGTCTGAGAGTGGTTGTGAGAGAGAGCTGCCAAG 1440

QY 1381 TTTCAACAGAGAACAGAAACCTTGTCTGAGAGTGGTTGTGAGAGAGAGCTGCCAAG 1440

DB 1381 TTTCAACAGAGAACAGAAACCTTGTCTGAGAGTGGTTGTGAGAGAGAGCTGCCAAG 1440

QY 1441 CCGCTTGTTCAGAGAAATTTCTTCTGTTATGCTTATGCTGATTAACATATGAGG 1500

DB 1441 CCGCTTGTTCAGAGAAATTTCTTCTGTTATGCTTATGCTGATTAACATATGAGG 1500

QY 1441 CCGCTTGTTCAGAGAAATTTCTTCTGTTATGCTTATGCTGATTAACATATGAGG 1500

DB 1441 CCGCTTGTTCAGAGAAATTTCTTCTGTTATGCTTATGCTGATTAACATATGAGG 1500

QY 1501 AGCTTAATGAGATTTAAGCAAGTATCATGAGCTGACGCTGACGCTGTATTTTCT 1560

DB 1501 AGCTTAATGAGATTTAAGCAAGTATCATGAGCTGACGCTGACGCTGTATTTTCT 1560

```

Db 1501 AGCTTACTATGCACTTTTAAAGCAAGATCCAGACGCTGCGAGCCGCGATCTTTTCT 1560
QY 1561 GGGGTGAGAGATCGCTAGAGTAAGATTTCTCTAAATTTATTTTGTCTTACTTTATT 1620
Db 1561 GGGGTGAGAGATCGCTAGAGTAAGATTTCTCTAAATTTATTTTGTCTTACTTTATT 1620
QY 1621 GCAGATGCTTCTTGTGCGGGGTGGGGGGTTTATTTGCTCCCAATGCTTTTGTATATCC 1680
Db 1621 GCAGATGCTTCTTGTGCGGGGTGGGGGGTTTATTTGCTCCCAATGCTTTTGTATATCC 1680
QY 1681 GGTGCTGTCTCTTATGTTGTCAGAGTGATGCTTCTGSCATTTTAAATTTGCTGTTGCC 1740
Db 1681 GGTGCTGTCTCTTATGTTGTCAGAGTGATGCTTCTGSCATTTTAAATTTGCTGTTGCC 1740
QY 1741 TTCCAGTTGGGAGATCATTTTACATMAACCGAAGATTCGGGATGATGATCTTCT 1800
Db 1741 TTCCAGTTGGGAGATCATTTTACATMAACCGAAGATTCGGGATGATGATCTTCT 1800
QY 1801 CAGTACTTTAATCATGTCGCTCTGCAACTTTTCTATCTGAGCCATCTTATCAACCAATC 1860
Db 1801 CAGTACTTTAATCATGTCGCTCTGCAACTTTTCTATCTGAGCCATCTTATCAACCAATC 1860
QY 1861 CTCTCAACCTCATTTTCAAGAAGTACAGAGCGCGGCTTTAACTGCTGCTCGCAAG 1920
Db 1861 CTCTCAACCTCATTTTCAAGAAGTACAGAGCGCGGCTTTAACTGCTGCTCGCAAG 1920
QY 1921 AAGTCAGAGCCGAGAGGCTTCCACAGAGCAGAGACACTGCGGGGAGATTTGACGGGAC 1980
Db 1921 AAGTCAGAGCCGAGAGGCTTCCACAGAGCAGAGACACTGCGGGGAGATTTGACGGGAC 1980
QY 1981 ACTGAGAGAGACACGGTGGCTTACACCCGACAGAGCGCTTAACTGTAAGACGATGGATTA 2040
Db 1981 ACTGAGAGAGACACGGTGGCTTACACCCGACAGAGCGCTTAACTGTAAGACGATGGATTA 2040

```

## RESULT 2

AL137000

## LOCUS

DEFINITION

Human DNA sequence from clone RP11-203116 on chromosome 13 contains the gene for KIAA0970 protein, COX7CPI (cytochrome c oxidase subunit VIc pseudogene 1), a novel pseudogene, the GPR38 (G protein-coupled receptor 38) gene, ESTs, STSs, GSSs and a Cpg island, complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL137000 163284 bp DNA linear PRI 18-DEC-2000  
 Human DNA sequence from clone RP11-203116 on chromosome 13 contains the gene for KIAA0970 protein, COX7CPI (cytochrome c oxidase subunit VIc pseudogene 1), a novel pseudogene, the GPR38 (G protein-coupled receptor 38) gene, ESTs, STSs, GSSs and a Cpg island, complete sequence.  
 AL137000 6 GI:9944121  
 HUG; COX7CPI; Cpg island; cytochrome c oxidase; G protein-coupled receptor; GPR38; KIAA0970.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 Wall, M.  
 Direct Submission  
 Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, UK  
 requests: clonerequests@sanger.ac.uk  
 On Aug 29, 2000 this sequence version replaced gi:9926419.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 The following abbreviations are used to associate primary accessions on the WormBase database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormbase This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at

## FEATURES

source

http://www.sanger.ac.uk/HGP/Chr13  
 IMPORTANT: This sequence is not the entire insert of clone RP11-203116. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
 The true left end of clone RP11-452110 is at 1 in this sequence.  
 The true right end of clone RP11-103118 is at 163284 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous there is an annotation using the 'unsure' feature key. RP11-203116 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACE3.6.  
 location/Qualifiers  
 1. 163284  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="13"  
 /clone="RP11-203116"  
 /clone\_11b="RP11-11.1"  
 3. 121  
 /note="MER20 repeat: matches 4. .126 of consensus"  
 152. .261  
 /note="MIR repeat: matches 20. .137 of consensus"  
 262. .554  
 /note="AluX repeat: matches 1. .293 of consensus"  
 555. .698  
 /note="MIR repeat: matches 137. .255 of consensus"  
 1012. .1291  
 /note="AluJ repeat: matches 7. .291 of consensus"  
 1320. .1425  
 /note="MER94 repeat: matches 4. .114 of consensus"  
 1591. .1804  
 /note="L1P8 repeat: matches 5945. .6158 of consensus"  
 3240. .3621  
 /note="L2 repeat: matches 2310. .2691 of consensus"  
 3824. .4106  
 /note="AluJ repeat: matches 1. .301 of consensus"  
 4975. .5042  
 /note="L2 repeat: matches 2639. .2705 of consensus"  
 6113. .6408  
 /note="AluX repeat: matches 1. .295 of consensus"  
 7646. .7732  
 /note="AluSc repeat: matches 1. .85 of consensus"  
 9887. .10191  
 /note="AluX repeat: matches 1. .306 of consensus"  
 12201. .12247  
 /note="L2 repeat: matches 2704. .2750 of consensus"  
 12472. .12738  
 /note="L2 repeat: matches 1304. .1530 of consensus"  
 12739. .12874  
 /note="L2 repeat: matches 2. .133 of consensus"  
 12875. .12887  
 /note="L2 repeat: matches 1530. .1542 of consensus"  
 12922. .13227  
 /note="L2 repeat: matches 5871. .6183 of consensus"  
 13303. .13348  
 /note="MIR repeat: matches 1. .46 of consensus"  
 13533. .13623  
 /note="MIR repeat: matches 223. .315 of consensus"  
 13657. .13865  
 /note="MIR repeat: matches 307. .533 of consensus"  
 13890. .14150  
 /note="MIR repeat: matches 195. .456 of consensus"  
 14151. .14450  
 /note="AluJ repeat: matches 1. .307 of consensus"  
 14501. .14729  
 /note="L1M4C repeat: matches 588. .813 of consensus"  
 14730. .15006  
 /note="AluJ repeat: matches 1. .285 of consensus"  
 15007. .15361

repeat_region	/note="L1M4c repeat: matches 813. .1178 of consensus" 15389. .15437 /note="L1M4c repeat: matches 1159. .1207 of consensus"
repeat_region	15438. .16412 /note="HERVL repeat: matches 2891. .3879 of consensus"
repeat_region	16413. .16616 /note="MLT2B repeat: matches 1. .216 of consensus"
repeat_region	17092. .17141 /note="MERSB repeat: matches 27. .79 of consensus"
repeat_region	17416. .17647 /note="MTR repeat: matches 17. .261 of consensus"
repeat_region	19402. .19499 /note="HY3 repeat: matches 2. .99 of consensus"
repeat_region	20730. .21163 /note="L2 repeat: matches 2168. .2609 of consensus"
repeat_region	25800. .25825 /note="13 copies 2 mer tt 92% conserved"
repeat_region	26041. .26131 /note="FLAM_C repeat: matches 17. .107 of consensus"
repeat_region	26144. .26249 /note="53 copies 2 mer tt 64% conserved"
repeat_region	27029. .27259 /note="MIR repeat: matches 2. .251 of consensus"
repeat_region	27422. .27788 /note="L2 repeat: matches 1879. .2259 of consensus"
repeat_region	27863. .28239 /note="L2 repeat: matches 1278. .1734 of consensus"
repeat_region	28255. .28571 /note="ALVJ repeat: matches 1. .310 of consensus"
repeat_region	28729. .29128 /note="L1M63 repeat: matches 4905. .5326 of consensus"
repeat_region	29170. .29413 /note="L1M63 repeat: matches 5461. .5715 of consensus"
repeat_region	29448. .29681 /note="L2 repeat: matches 1594. .1828 of consensus"
repeat_region	29683. .29886 /note="L1M6 repeat: matches 5527. .5734 of consensus"
repeat_region	29888. .30323 /note="L2 repeat: matches 2129. .2596 of consensus"
repeat_region	30324. .30620 /note="ALUtb repeat: matches 1. .299 of consensus"
repeat_region	30621. .30780 /note="L2 repeat: matches 2596. .2749 of consensus"
repeat_region	30784. .30813 /note="15 copies 2 mer aa 86% conserved"
repeat_region	30814. .30947 /note="L1M63 repeat: matches 5724. .5874 of consensus"
repeat_region	33134. .33214 /note="MIR repeat: matches 175. .262 of consensus"
repeat_region	35454. .35753 /note="ALUx repeat: matches 1. .300 of consensus"
misc_feature	36245. .36743 /note="match: GSS: Em:AO593229"
repeat_region	36656. .36843 /note="MIR repeat: matches 14. .237 of consensus"
misc_feature	38014. .38191 /note="match: STS: Em:G04228"
repeat_region	38279. .38416 /note="MERSA repeat: matches 7. .152 of consensus"
repeat_region	39603. .39708 /note="MIR repeat: matches 34. .144 of consensus"
repeat_region	40037. .40103 /note="L2 repeat: matches 2664. .2731 of consensus"
repeat_region	40370. .40523 /note="L2 repeat: matches 2562. .2701 of consensus"
repeat_region	40524. .40810 /note="ALUx repeat: matches 2. .297 of consensus"
repeat_region	40811. .41080 /note="L2 repeat: matches 2242. .2562 of consensus"
repeat_region	41085. .41156 /note="L2 repeat: matches 2626. .2697 of consensus"
repeat_region	41549. .41871 /note="MERS2 repeat: matches 1. .343 of consensus"



Db	145068	CAGCTGGGCGCGCTGCGCTGATGCTGATGGGTCAACACCGCGCTACTTCTTCGCGCCCTTT	145122
QY	781	CTGGGCTCAGACATTCCTCTAATCGGGCTCATTCGGGCGGGAAGCTGTGAGCAAGCCGGCGCG	840
Db	145128	CTGTGCTTCACACATCTCTTACGGGCTCATTCGGGCGGGAAGCTGTGAGCAAGCCGGCGCG	145187
QY	841	CTGGGAGCCCGGCGCCTTCGGGCGGGAAGAGGCAACGGGAGACAAACGCTGCTG	900
Db	145188	CTGGAGGCGCGCGCCTTCGGGCGGGAAGAGGCAACGGGAGACAGCTGCGGCTCTG	145247
QY	901	CGTAAGTGAAGCCCGCTGTGTTCCAAAGACGCTGCTGCAAGTCGCGCCCGCGGGAGC	960
Db	145248	CGTAAGTGAAGCCCGCTGTGTTCCAAAGACGCTGCTGCAAGTCGCGCCCGCGGGAGC	145307
QY	961	GCGCAACGCTGGGCGCCCTTCCCTGCGCGCAAGCTGTGGGCGCGGCTTCAGCTGCC	1020
Db	145308	GCGCAACGCTGGGCTCCCTTCCCTGCGCGCAAGCTGTGGGCGCGGCTTCAGCTGCC	145367
QY	1021	TTTTCTATTTTGATTTCAAGCCTCCACCCGCGGTACTTCCATCCCGGAGAAACATG	1080
Db	145368	TTTTCTATTTTGATTTCAAGCCTCCACCCGCGGTACTTCCATCCCGGAGAAACATG	145427
QY	1081	TTCGTGTCGCCAGGAGCTCTGGGGGACCCGAGGGGCTTTGAGGGTGGATCCCGGATC	1140
Db	145428	TTCGTGTCGCCAGGAGCTCTGGGGGACCCGAGGGGCTTTGAGGGTGGATCCCGGATC	145487
QY	1141	CGATTCAGTAACCAAGCAGTGTCTTTTCCAGAGCCTCTGAGACAGAAAGAGAGTTGGTA	1200
Db	145488	CGATTCAGTAACCAAGCAGTGTCTTTTCCAGAGCCTCTGAGACAGAAAGAGAGTTGGTA	145547
QY	1201	TTCTTATCCCAACCACTGTTAATGCCAATAAGAGATCTTCAAGTGTCTTTGAGA	1260
Db	145548	TTCTTATCCCAACCACTGTTAATGCCAATAAGAGATCTTCAAGTGTCTTTGAGA	145607
QY	1261	AGACGAGGAGATTTCAATTAAGCTAAAAATTTTTTAATGTTAAGATGCTGAGGC	1320
Db	145608	AGACGAGGAGATTTCAATTAAGCTAAAAATTTTTTAATGTTAAGATGCTGAGGC	145667
QY	1321	TAAAGTAAACCTTGCTGCTGATCAAAAAGTAAGATTGTCAGACCTGTGTGAATTCTT	1380
Db	145668	TAAAGTAAACCTTGCTGCTGATCAAAAAGTAAGATTGTCAGACCTGTGTGAATTCTT	145727
QY	1381	TTCAACAGAGAACAGAAAATGTGCTCCGAGTGGGTTGTGGAAGAAAGCTGCCAAG	1440
Db	145728	TTCAACAGAGAACAGAAAATGTGCTCCGAGTGGGTTGTGGAAGAAAGCTGCCAAG	145787
QY	1441	CGGCTTGTTCAGAGAAATGTGCTTGTGTTAATGTCACGCTTGATAACATATGGG	1500
Db	145788	CGGCTTGTTCAGAGAAATGTGCTTGTGTTAATGTCACGCTTGATAACATATGGG	145847
QY	1501	AAGCTTCTATGCAATTTTAAAGCAAGTATCATGCAAGCTGCAAGCTGCTGCTATTTTTCT	1560
Db	145848	AAGCTTCTATGCAATTTTAAAGCAAGTATCATGCAAGCTGCTGCTGCTATTTTTCT	145907
QY	1561	GGGGTAGGATCTGCGCTAGAGTAAGTTTCTCTAATTTTGGCTGCTGCTGCTGCT	1620
Db	145908	GGGGTAGGATCTGCGCTAGAGTAAGTTTCTCTAATTTTGGCTGCTGCTGCTGCTGCT	145967
QY	1621	GCAAGATGTTCTTGTGCGGGGTGGGGGTTTATTTGCTTCCAAATGCTTTGTAATCC	1680
Db	145968	GCAAGATGTTCTTGTGCGGGGTGGGGGTTTATTTGCTTCCAAATGCTTTGTAATCC	146027
QY	1681	GGTGCCTGCTTATGATGTGAGTGGTGGTTCGCAATTTAATAATTTGCTGTGGCC	1740
Db	146028	GGTGCCTGCTTATGATGTGAGTGGTGGTTCGCAATTTAATAATTTGCTGTGGCC	146087
QY	1741	TTTCCAGTTTGGAGAAATCATTTATCATAAACGGAAGATTGCGGATGATTAATCTCT	1800
Db	146088	TTTCCAGTTTGGAGAAATCATTTATCATAAACGGAAGATTGCGGATGATTAATCTCT	146147
QY	1801	CAGTCTTAAACATGCTGCTTGCACTTTTCTATCTGAGCGATCTATCAACCAATC	1860

Db 146148 CAGTACTTTAATCATGTCGTCTGCAACTTTTCTATCTGAGCGCATTAATCAACCAATC 146207

QY 1861 CTTTACAACTCTATTTTCAAAAGATACAGAGCGCGCTTTTAACTGCTCTGCAAG 1920

Db 146208 CTTTACAACTCTATTTTCAAAAGATACAGAGCGCGCTTTTAACTGCTCTGCAAG 146267

QY 1921 AAGTCCAGGCCGAGAGGCTTCCACAGAGCAGGGACACTGCGGGGGAAGTTGCAGGGGAC 1980

Db 146268 AAGTCCAGGCCGAGAGGCTTCCACAGAGCAGGGACACTGCGGGGGAAGTTGCAGGGGAC 146327

QY 1981 ACTGAGAGAGACAGGTGTGGGCTACACCGAGACAAGCGCTAACGTGGAAGAGATGGGATTA 2040

Db 146328 ACTGAGAGAGACAGGTGTGGGCTACACCGAGACAAGCGCTAACGTGGAAGAGATGGGATTA 146387

```

RESULT 3
AL596304
LOCUS AL596304 341560 bp DNA linear HTG 09-AUG-2001
DEFINITION Homo sapiens chromosome 13 clone RP11-804, *** SEQUENCING IN
ACCESSION AL596304
VERSION AL596304.3 GI:15147703
KEYWORDS HTG, HTGS_PHASE1, HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 341560)
AUTHORS Clark,G.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clone.requests@sanger.ac.uk
On Aug 10, 2001 this sequence version replaced gi:15131387.

COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BA804
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 149716 bases at least Q40
Consensus quality: 150039 bases at least Q30
Consensus quality: 150247 bases at least Q20
Insert size: 341360; sum-of-contigs
Insert size: 157238; 2.0% error; agarose-fp
Quality coverage: 3.92x in Q20 bases; sum-of-contigs quality
coverage: 8.77x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 146853: contig of 146853 bp in length
* 146854 146953: gap of 100 bp
* 146954 338660: contig of 191707 bp in length
* 338661 338760: gap of 100 bp
* 338761 341560: contig of 2800 bp in length.
*
Location/Qualifiers
1..341560
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-804"
/clone_1lb="RP11-11.1"
FEATURES
source

```







Db	440	TGCGCAGCCGTCCGCACGCCGCCGCTTTTGGTCACCACCGGACCGCGTGCCGCGGCTCATCGCT	499
OY	481	GTCCTCGGGCCGTGGGCGGTGCTCTGCGCGGCCCCTTCTGTCTCTGATGGGCGTCGAG	540
Db	500	GTCCTCTGGGCCCTTGCGGCTGCTCTCTGCGCGGCTCTTCTTGTGTGGAGCGTCGAG	559
OY	541	CAGAACCCCGGCATCTCCGTAGTCCCCGGGCTTCMATGGACCGCGGAGATCGCTCTCG	600
Db	560	CAGAGACCCCGGCATCTTCNCGTAGTCCCGGGGCGCTCATGGACCGGNGCGATTNNCGCTCTG	619
OY	601	CCCTCGGCGCTCGGCGGCGGCGCTCTCTGAGCTCTCGCGGAGCGCACCGCGCTCCCGCGCTCG	660
Db	620	CTTTTGCTGNMGNGCGCGCTTTTGGG-----TCTGCGGCGCGCACCGCGTCCCGCTTCG	674
OY	661	GGGCGCCGAGACGCGG	675
Db	675	GGGCGCCGAGACGCGG	689
RESULT 7			
HSAS42408		692 bp	DNA linear PRI 18-JUL-2002
LOCUS			
DEFINITION	Homo sapiens genomic sequence surrounding NotI site, clone NR1-PB24C.		
ACCESSION	AJ342408		
VERSION	AJ342408.1	GI:15886883	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 692)		
	Kutuemov,A.S., Gizaullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M., Podomski,I.R.M., Matuchkin,Y.G., Gyanchandani,A., Muravenko,O.V., Levitsky,V.G., Kolchanov,N.A., Protodopov,A.I., Kashba,V.I., Kiselev,L.L., Wasserman,W., Mahstedt,C. and Zharovsky,E.R.		
TITLE	NotI flanking sequences: a tool for gene discovery and verification of the human genome		
JOURNAL	Nucleic Acids Res.	30 (14),	3163-3170 (2002)
MEDLINE	22131767		
PUBMED	12136098		
REFERENCE	2 (bases 1 to 692)		
AUTHORS	Zabarovskiy,E.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-MAY-2001) Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden		
FEATURES			
source	Location/Qualifiers		
	1..692		/organism="Homo sapiens"
	/db_xref="taxon:9606"		
	/clone="NR1-PB24C"		
BASE COUNT	92 a 239 c 219 g	140 t	2 others
ORIGIN			
Query Match	28.0%; Score 572.2;	DB 9;	Length 692;
Best Local Similarity	.91.5%;	Pred. No. 8,1e-79;	
Matches	616; Conservative	0; Mismatches	55; Indels
			2; Gaps
OY	1	ATGGGCAAGCCCTTGGAACGACGACGAGCGCCCGAAGGGGGCGCGGAGACCGCGTGAGCCC	60
Db	20	ATGGGCAAGCCCTTGGAACGAGCGACGAGCGCCCGAAGGGGGCGCGGAGACCGCGTGAGCCC	79
OY	61	GCGGCGCGCCCTTGGAAGGAGGCGGCGGCGTCCGCTTCCCTGAGGGCGCGTGCGCG	120
Db	80	GCGGCGCGCCCTTGGAAGGAGGCGGCGGCGTCCGCTTCCCTGAGGGCGCGTGCGCG	139
OY	121	GTCACCGGCTGTGCTGCTTGCCTGTTCTGTCGTCGGGGGTGACGCGCAAATGTGTACCGTG	180
Db	140	GTCACCGGCTGTGCTGCTGCTGTTCTGTCGTCGTCGTCGTCGTCGTCGTAACCGTG	199
OY	181	ATGTCGATCGGGCGCTTACCGGAGATGCGGACACACAACTTGTACTTCGGGACGATG	240

Db	200	ATGCTATGGGGGCTACCGGGACATGGGACCAACCAACTTTGACTGGGACACATG	259
QY	241	GCCGTATCCGACCTACTCATCTCTGCTGGGACTGCCGTTGACCTTACCCTCTGGCCG	300
Db	260	GCCGTATCCGACCTACTCATCTCTGCTGGGACTGCCGTTGACCTTACCCTCTGGCCG	319
QY	301	TGCGGGCCCTGGGGTGTTCGGGGCCGCGTCTCTGCGGCTCTCCCTCTACTGTGGGCGAGGC	360
Db	320	TGCGGGCCCTGGGGTGTTCGGGGCCGCGATGCTATGCCCTTACTGTGGGCGAGGGC	379
QY	361	TGCACCTACGCCACGCTGCTGACATGACCGGCTAGCGTGAAGCGCTACCTGGCCATC	420
Db	380	TGCACCTACGCCACGCTGCTGACATGACCGGCTAGCGTATAGGCTACCTTGGCATC	439
QY	421	TGCGGCGGCGCTCCGGCCCGCGCTTTGGTGCACCCGGGCGCGCTCCGCTCATCGCT	480
Db	440	TGCGGCGGCGCTCCGGCCCGGACACGCGTATGGTACCCGGCGCGGATACGCTCATCGCT	499
QY	481	GTGCTCTGGGGCGGTGGGCGCTGCTCTCTGCGCGGTGCCCTTTGTTCTGCTGGGCGGCGAG	540
Db	500	GTGCTCTGGGGCGGTGGGCGCTGCTCTATATGACATCCTCTTTTGTCTGGGCGGCGTCTAG	559
QY	541	CAGGACCCCGGCATCTCCGTAGTCCCGGCGCTCAATGGACCGCGCGAGATCGCTCTCG	600
Db	560	CATGACACCTGTATCTCGTACTCTCGGGCCCATATGCAACCGAGTAACTACTTC	619
QY	601	CGT--CTGGCCCTGGGCGGCGCTCTCTGGCTCTCGCGGGCGGACCGCGTCCCGCGGT	658
Db	620	GCTTATATGACATATGTGACGACTCTGTTGATATCGCGGGTGCAACGACATACATACCGT	679
QY	659	CGGGGCGCGGAGAC	671
Db	680	CAGGCGCGCATAC	692
RESULT 8			
HSA339459			
LOCUS		615 bp	DNA linear PRI 18-JUL-2002
DEFINITION		Homo sapiens genomic sequence surrounding Not1 site, clone	
ACCESSION		NR1-WC14C.	
VERSION		AJ339459	
KEYWORDS		AJ339459.1 GI:15883877	
ORGANISM			
SOURCE		Homo sapiens.	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		1 (bases 1 to 615)	
JOURNAL		Kurshenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,	
MEDLINE		Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,	
PUBMED		Levitelsky, V.G., Kolchanov, N.A., Protolopov, A.I., Kashuba, V.I.,	
REFERENCE		Kisselev, L.L., Wasserman, W., Wahlstedt, C., and Zabarovsky, E.R.	
AUTHORS		Not1 flanking sequences: a tool for gene discovery and verification	
TITLE		of the human genome	
JOURNAL		Nucleic Acids Res. 30 (14), 3163-3170 (2002)	
MEDLINE		22331767	
PUBMED		12136098	
REFERENCE		2 (bases 1 to 615)	
AUTHORS		Zabarovsky, E.R.	
TITLE		Direct Submission	
JOURNAL		Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,	
FEATURES		Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,	
SOURCE		Sweden	
LOCATION/Qualifiers		1..615	
ORGANISM="Homo sapiens"		/db_xref="taxon:9606"	
clone="NR1-WC14C"		/clone="NR1-WC14C"	
61 a	224 c	204 g	121 t 5 others
Query Match		27.3%	Score 557.4; DB 9; Length 615;



TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Mus musculus, clone RP24-199E5

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campiolo, A., Chang, J., Chazaro, B., Chopel, Y., Colangelo, M., Collipins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gadyas, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iley, I., Johnson, R., Jones, C., Kamat, A., Karczas, A., Kells, C., Lacroque, K., Lamareas, R., Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Milova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nobbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunhnan, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupp, R., Seaman, S., Severy, P., Spencer, B., Stange-Rothmann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talmes, J., Tefaye, S., Theodore, J., Toplak, K., Travers, M., Travis, N., Trigilio, J., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE	REFERENCE	AUTHORS
1. The effect of the concentration of the solution on the rate of the reaction.	1. The effect of the concentration of the solution on the rate of the reaction.	1. The effect of the concentration of the solution on the rate of the reaction.
2. The effect of the temperature on the rate of the reaction.	2. The effect of the temperature on the rate of the reaction.	2. The effect of the temperature on the rate of the reaction.
3. The effect of the catalyst on the rate of the reaction.	3. The effect of the catalyst on the rate of the reaction.	3. The effect of the catalyst on the rate of the reaction.
4. The effect of the surface area of the solid reactant on the rate of the reaction.	4. The effect of the surface area of the solid reactant on the rate of the reaction.	4. The effect of the surface area of the solid reactant on the rate of the reaction.
5. The effect of the pressure on the rate of the reaction.	5. The effect of the pressure on the rate of the reaction.	5. The effect of the pressure on the rate of the reaction.
6. The effect of the volume of the gas reactant on the rate of the reaction.	6. The effect of the volume of the gas reactant on the rate of the reaction.	6. The effect of the volume of the gas reactant on the rate of the reaction.
7. The effect of the concentration of the reactants on the rate of the reaction.	7. The effect of the concentration of the reactants on the rate of the reaction.	7. The effect of the concentration of the reactants on the rate of the reaction.
8. The effect of the concentration of the products on the rate of the reaction.	8. The effect of the concentration of the products on the rate of the reaction.	8. The effect of the concentration of the products on the rate of the reaction.
9. The effect of the concentration of the catalyst on the rate of the reaction.	9. The effect of the concentration of the catalyst on the rate of the reaction.	9. The effect of the concentration of the catalyst on the rate of the reaction.
10. The effect of the concentration of the solvent on the rate of the reaction.	10. The effect of the concentration of the solvent on the rate of the reaction.	10. The effect of the concentration of the solvent on the rate of the reaction.

Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 172168)

[illegible]

TITLE  
JOURNAL  
COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA  
 Submitted (25-Apr-2002) Whitehead Institute/MIT Center for Genome  
 Research  
 On Apr 25, 2002 this sequence version replaced gi:18482076.  
 All repeats were identified using RepeatMasker:  
 Smith, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIMR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information

```
----- Summary Statistics -----
Sequencing vector: Plasmid; n/a, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 165510 bases at least Q40
Consensus quality: 167875 bases at least Q30
Consensus quality: 168675 bases at least Q20
```

Insert size: 176000; agarose-fp  
Insert size: 169868; sum-of-contigs  
Quality coverage: 8.4 in Q20 bases; agarose-fp  
Quality coverage: 8.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

1	291:	contig of 291 bp	in length
292	391:	gap of 100 bp	
392	1443:	contig of 1052 bp	in length
1444	1543:	gap of 100 bp	
1544	2462:	contig of 919 bp	in length
2463	2562:	gap of 100 bp	
2563	3236:	contig of 674 bp	in length
3237	3333:	gap of 100 bp	
3337	4351:	contig of 1017 bp	in length
4354	4453:	gap of 100 bp	
4454	5453:	contig of 1000 bp	in length
5454	5553:	gap of 100 bp	
5554	6971:	contig of 1418 bp	in length
6972	7071:	gap of 100 bp	
7072	8137:	contig of 1066 bp	in length
8138	8237:	gap of 100 bp	
8238	10559:	contig of 2332 bp	in length
10570	10669:	gap of 100 bp	
10670	13572:	contig of 2903 bp	in length
13573	13672:	gap of 100 bp	
13673	16655:	contig of 2983 bp	in length
16566	16755:	gap of 100 bp	
16756	20655:	contig of 3880 bp	in length
20636	20735:	gap of 100 bp	
20736	22907:	contig of 2172 bp	in length
22908	23007:	gap of 100 bp	
23008	26263:	contig of 3356 bp	in length
26264	26363:	gap of 100 bp	
26364	31037:	contig of 4674 bp	in length
31038	31137:	gap of 100 bp	
31138	38064:	contig of 6627 bp	in length
38065	38164:	gap of 100 bp	
38165	48247:	contig of 10083 bp	in length
48248	48347:	gap of 100 bp	
48348	57922:	contig of 9575 bp	in length
57923	58022:	gap of 100 bp	
58023	71658:	contig of 13606 bp	in length
71629	71728:	gap of 100 bp	
71729	90376:	contig of 18648 bp	in length
90377	90476:	gap of 100 bp	
90477	107535:	contig of 17059 bp	in length
107536	107635:	gap of 100 bp	
131188	131187:	contig of 33552 bp	in length
131288	13287:	gap of 100 bp	
152395	152394:	contig of 2107 bp	in length
152395	152494:	gap of 100 bp	
152495	172168:	contig of 19674 bp	in length

**FEATURES**  
**SOURCE**

```
misc_feature      1.291
                 /note="assembly_fragment"
misc_feature      392.143
                 /note="assembly_fragment"
misc_feature      1544.2462
                 /note="assembly_fragment"
```



Query Match	15.0%;	Score 307;	DB 2;	Length 172168;
Best Local Similarity	70.2%;	Pred. No. 4.5e-38;		
Matches 530;	Conservative	0;	Mismatches 180;	Indels 45;
				Gaps 77;

Db	35277	TCATTGCTTCGCTTGGGTTTGCGCCTGTCTC-----TACCTTCTTCTTCTGATGG	35225
QY	533	GGCTCGAGAGAGACCCGGGATCTCCGTAGTCCCGGACCTCAATGGACCGCGGATCG	592
Db	35224	GGCTTCATCAAGATGTCCTCCCGGCTCTCTGCGGACCTGGATCAACATTTGGCAATCA	35165
QY	593	CGTCTTCGCTCTTCGCTCTGTCGCGCCCTCTGAGGCTTCGCGGCGCCACCGCCCTGCC	652
Db	35164	ACCCCTCGCTCCGCGCTCGACACTTCTCTTGGGTCTTGAAGACACCAACGCTGTCCC	35105
QY	653	CGCCGTGCGGGGCCGAGACCGCGAGAGCCCGGCGCTGTTCAGCCCGGAATGCCGCGCA	712
Db	35104	CGCAATATGGAACCAAGCGCTGAGAGC-----GTTCAAGCCGGAATGGCC-ATCGA	35055
QY	713	GCCCCGCGAGCTGGGGCGCTCGGTGTATGCTGTGGGTGATCAACCGGCTACTTCTTCC	772
Db	35054	GCCCCACTCAGTAGAGGCTCTGGGGGTCTGTGTGGGTACACACCGCTTA--TTTCC	34998
QY	773	TGCCCCCTTCTGTGCTCTGACATCCTGTACGAGGCTCATCGGACGGAGACTGTGAGCAGCC	832
Db	34997	TGCCCCCTTCTGTGCTCTGAGTCTCTTCTTCCGAGCTCTGTC-GGCGTTACCGTGAAGGCC	34939
QY	833	GGCGCGCGCTGCGAGGCGCCGCGCTCGGGGCGG	867
Db	34938	GGGTCCCGACCAAGGTGTCGGAAGAAAGGGCGG	34904
RESULT 11			
LOCUS	ARI68467	283 bp	DNA
DEFINITION	Sequence 4 from patent US 6287855.		linear
ACCESSION	ARI68467		
VERSION	ARI68467.1	GI:17904400	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 283)		
TITLE	Tan, C. and Sullivan, K.		
JOURNAL	Nucleic acid encoding rat galanin receptor (GALR2)		
FEATURES	Patent: US 6287855-A 4 11-SEP-2001;		
Source	Location/Qualifiers		
	1..283		
	/organism="unknown"		
BASE COUNT	27 a	116 c	84 g
ORIGIN		56 t	
Query Match 13.9%; Score 283; DB 6; Length 283;			
Bee Local Similarity 100.0%; Pred. No. 3.9e-34;			
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	206	TGCGAGCACCAACCAACTTGTACTGTGGGAGCATGGCGGTGTCGACCTACTCATCTTCGC	265
Db	1	TGCGAGCACCAACCAACTTGTACTGTGGGAGCATGGCGGTGTCGACCTACTCATCTTCGC	60
QY	266	TGGGCTGCGCTTGCACCTGTACCGGCTCTGGCGGCTCGGCGGCGCTGGGTTTCGGGCGCG	325
Db	61	TGGGCTGCGCTTGCACCTGTACCGGCTCTGGCGGCTCGGCGGCGCTGGGTTTCGGGCGCG	120
QY	326	TGCTCTGCGGCTGCTCCCTCTACGTGTGGGGAAGGCTGCACTTACGCCACGCTGTGACACA	385
Db	121	TGCTCTGCGGCTGCTCCCTCTACGTGTGGGGAAGGCTGCACTTACGCCACGCTGTGACACA	180
QY	386	TGACGCGGCTACGCGTGAAGGCTTACCTGGCAATCTGCGCGCGCTTCGCGCGCGGCTCT	445
Db	181	TGACGCGGCTACGCGTGAAGGCTTACCTGGCAATCTGCGCGCGCTTCGCGCGCGGCTCT	240
QY	446	TGGTCAACCGGCGCGCGCTTCGCGCGGCTCATGCGTGTGCTTG	488
Db	241	TGGTCAACCGGCGCGCGCTTCGCGCGGCTCATGCGTGTGCTTG	283
RESULT 12			



AR182284  
 LOCUS AR182284 283 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 4 from patent US 6337206.  
 ACCESSION AR182284  
 VERSION AR182284.1 GI:20225200  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 283)  
 AUTHORS Tan, C. and Kolakowski, L.F., Jr.  
 TITLE Nucleic acid encoding mouse galanin receptor (GALR2)  
 JOURNAL Patent: US 6337206-A 4 08-JAN-2002;  
 FEATURES  
 Source  
 1..283  
 Location/Qualifiers  
 BASE COUNT 27 a 116 c 84 g 56 t  
 ORIGIN  
 Query Match 13.9%; Score 283; DB 6; Length 283;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-34;  
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 206 TGGGACCAACCACTGTAAGTGGGAGCATAGCCGTGTCGAACTTCACTCTGC 265  
 Db 1 TGGGACCAACCACTGTAAGTGGGAGCATAGCCGTGTCGAACTTCACTCTGC 60  
 QY 266 TGGGAGTCCGTTTCGACCTGTACCGGCTGCGGCTCGGAGCCCTGGTGTGGGCGG 325  
 Db 61 TGGGAGTCCGTTTCGACCTGTACCGGCTGCGGCTCGGAGCCCTGGTGTGGGCGG 120  
 QY 326 TGTCTGCGGCTGTGTCCTCTCAAGTGGGCGAGGGCTGACCTACGCGAGCTGTGACA 385  
 Db 121 TGTCTGCGGCTGTGTCCTCTCAAGTGGGCGAGGGCTGACCTACGCGAGCTGTGACA 180  
 QY 386 TGACCGGCTGAGGCTGAGGCGCTACCTGGCCATGCGCGCCGCGGCGCGGCTCT 445  
 Db 181 TGACCGGCTGAGGCTGAGGCGCTACCTGGCCATGCGCGCCGCGGCGCGGCTCT 240  
 QY 446 TGGTCAACCCGCGCGCGCTGCGGCTCATCGCTGTGCTCTG 488  
 Db 241 TGGTCAACCCGCGCGCGCTGCGGCTCATCGCTGTGCTCTG 283  
 RESULT 13  
 HSA341295 573 bp DNA linear PRI 18-JUL-2002  
 LOCUS HSA341295 Homo sapiens genomic sequence surrounding Nct1 site, clone  
 DEFINITION N5-KN11C.  
 ACCESSION AJ341295  
 VERSION AJ341295.1 GI:15885713  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 573)  
 AUTHORS Kutsenko, A.S., Gizaullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M.,  
 Podowski, R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V.,  
 Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I.,  
 Kiselev, L.L., Wasserman, W., Wahlstedt, C., and Zabarovsky, E.R.  
 TITLE Nct1 flanking sequences: a tool for gene discovery and verification  
 of the human genome  
 JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)  
 MEDLINE 2213167  
 PUBMED 12136098  
 REFERENCE 2 (bases 1 to 573)  
 AUTHORS Zabarovsky, E.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2001) Microbiology and Tumorigenology Centre,  
 Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,  
 Sweden  
 Location/Qualifiers

source 1..573  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="N5-KN11C"  
 BASE COUNT 64 a 185 c 195 g 129 t  
 ORIGIN  
 Query Match 13.1%; Score 266.8; DB 9; Length 573;  
 Best Local Similarity 98.9%; Pred. No. 1.2e-31;  
 Matches 279; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 ATGGGAGCCCTTGGAAAGCGGACGAGCGCCCGAGAGGGGCGGAGACCGCGTGGCC 60  
 Db 20 ATGGGAGCCCTTGGAAAGCGGACGAGCGCCCGAGAGGGGCGGAGACCGCGTGGCC 79  
 QY 61 GGGGCGGCGCTTGGAGAGGAGCGCCCGGTGTCGCGCCCTTCCCTGGGGGCGTGGTCC 120  
 Db 80 GGGGCGGCGCTTGGAGAGGAGCGCCCGGTGTCGCGCCCTTCCCTGGGGGCGTGGTCC 139  
 QY 121 GTGACCGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 180  
 Db 140 GTGACCGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 199  
 QY 181 ATGCTGATGGGCGCTACCGGAGATGCGGACCAACCACTTGTACTGTGGGAGCATG 240  
 Db 200 ATGCTGATGGGCGCTACCGGAGATGCGGACCAACCACTTGTACTGTGGGAGCATG 259  
 QY 241 GCGGTGCGGACCTACTCAATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 282  
 Db 260 GCGGTGCGGACCTACTCAATCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 300  
 RESULT 14  
 AR156351 1063 bp DNA linear PAT 08-AUG-2001  
 LOCUS AR156351  
 DEFINITION Sequence 1 from patent US 6242199.  
 ACCESSION AR156351  
 VERSION AR156351.1 GI:15125055  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1063)  
 AUTHORS Ploeg, L.H.T., Feighner, S.D., Howard, A.D., Pong, S.-S., and Van Der  
 Ploeg, L.H.T.  
 TITLE Assays for growth hormone secretagogue receptors  
 JOURNAL Patent: US 6242199-A 1 05-JUN-2001;  
 FEATURES  
 Source  
 1..1063  
 Location/Qualifiers  
 BASE COUNT 166 a 339 c 285 g 253 t  
 ORIGIN  
 Query Match 11.5%; Score 234.8; DB 6; Length 1063;  
 Best Local Similarity 69.3%; Pred. No. 9.6e-27;  
 Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
 QY 92 GCGCTTCCCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 151  
 Db 69 GCGCTTCCCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128  
 QY 152 TGGGAGTATGCGGAGCAAGTGTGACCGGATGTGTGTCGAGGCGGCTACCGGAGATGCGGA 211  
 Db 129 TGGGAGTATGCGGAGCAAGTGTGACCGGATGTGTGTCGAGGCGGCTACCGGAGATGCGGA 188  
 QY 212 CCACCAACAATTGTAAGTGTGAGGAGATGCGGCTGTGCGACCTATCTCTGCTGGGCG 271  
 Db 189 CCACCAACAATTGTAAGTGTGAGGAGATGCGGCTGTGCGACCTATCTCTGCTGGCA 248  
 QY 272 TGGGTTGACCTGTACCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 331  
 Db 249 TGGGTTGACCTGTACCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 308

QY 332 GCGGCTGTCCTCTACGTGGGCGAGGCTGCACCTACGCGACGCTGTGCATGACCG 391  
 Db 309 GCAACTCTTCCAGTGTGTAGAGAGAGCTGCACCTTACGCGACAGTGTCTACCTACCG 368  
 QY 392 CGCTGAGCGTGCAGCGCTACCTGCGCATCTGCGCGCTCGCGCGCTGCTGTGCA 451  
 Db 369 CGCTGAGCGTGCAGCGCTACCTGCGCATCTGCGCGCTCGCGCGCAAGTGTGCA 428  
 QY 452 CCGGCGCGCGCTCGCGCGCTGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCG 511  
 Db 429 CCAAGGCGCGCGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG 488  
 QY 512 GTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553  
 Db 489 GAGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530  
 RESULT 15  
 SSU60178  
 LOCUS SSU60178 1101 bp mRNA linear MAM 23-AUG-1996  
 DEFINITION Sus scrofa growth hormone secretagogue receptor type 1a mRNA,  
 complete cds.  
 ACCESSION U60178  
 VERSION U60178.1 GI:1504144  
 KEYWORDS  
 SOURCE Sus scrofa.  
 ORGANISM Sus scrofa.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Howard,A.D., Feighner,S.D., Cully,D.F., Arena,J.P., Liberatore,P.A.,  
 Rosenblum,C.I., Hamelin,M., Hreniuk,D.L., Palyla,O.C., Anderson,J.,  
 Pares,P.S., Diaz,C., Chou,M., Liu,K.K., McKee,K.K., Pong,S.-S.,  
 Chaung,L.-Y., Eibrecht,A., Dashkevich,M., Heavens,R., Rigby,M.,  
 Strimachsingh,J.D.J.S., Dean,D.C., Melillo,D.G., Patchett,A.A.,  
 Nargund,R., Griffin,P.R., Demartino,J.A., Gupta,S.K.,  
 Schaeffer,J.M., Smith,R.G. and Van der Ploeg,L.H.T.  
 A receptor in pituitary and hypothalamus that functions in growth  
 hormone release  
 Science 273 (5277), 974-977 (1996)  
 JOURNAL  
 MEDLINE  
 PUBMED 96337998  
 8688086  
 TITLE 2 (bases 1 to 1101)  
 JOURNAL Howard,A.D., Feighner,S.D., Cully,D.F., Arena,J.P., Liberatore,P.A.,  
 Rosenblum,C.I., Hamelin,M., Hreniuk,D.L., Palyla,O.C., Anderson,J.,  
 Pares,P.S., Diaz,C., Chou,M., Liu,K.K., McKee,K.K., Pong,S.-S.,  
 Chaung,L.-Y., Eibrecht,A., Dashkevich,M., Heavens,R., Rigby,M.,  
 Strimachsingh,J.D.J.S., Dean,D.C., Melillo,D.G., Patchett,A.A.,  
 Nargund,R., Griffin,P.R., Demartino,J.A., Gupta,S.K.,  
 Schaeffer,J.M., Smith,R.G. and Van der Ploeg,L.H.T.  
 Direct Submission  
 Submitted (06-JUN-1996) Genetics & Molecular Biology, Merck  
 Research Laboratories, P.O. Box 2000, Building RY-80Y-265, Rahway,  
 NJ 07065, USA  
 FEATURES  
 SOURCE  
 CDS  
 1. 1101  
 /location="Qualifiers  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /tissue\_type="anterior pituitary"  
 1. 1101  
 /function="growth hormone release"  
 /note="g-protein-alpha1 coupled receptor of pituitary and  
 hypothalamus; receptor for growth hormone releasing  
 peptides (e.g. GHRP-6, GHRP-2) and non-peptide, low  
 molecular weight secretagogues (e.g. L-692,429, MK-0677)"  
 /codon\_start=1  
 /product="growth hormone secretagogue receptor type 1a"  
 /protein\_id="AAC48630.1"  
 /db\_xref="GI:1504145"  
 /translation="WMNATPSEBPGNLTIPIGWADAPENDSTVBEILPLPPTPLA  
 GVTATCAALFVAGLIGNLLTMIIVSRFRMRTTNLYISSNAFSLILFLCPLDLFR  
 LMQYPMNIGNLKLFQFVSESCYATVLTITLALSVETFAICFPLRAKAVITGRV  
 KLIVLIVAWAFCSAGPIFVLVGVHDNGTDRDNECRATEFAVRSGLLTVMWVSS

VPEFLPVLCITLVYSLIGRKIMRRKREAAVSSLRDQNHQYKMLAVVFAFLICM  
 LPEHVGRIYLFPSKSLPSGVEIAQISQYCNIVSFVLFYLSAANPILNIMSKINAV  
 FKLIGFPPFQSRKSLTKDESSRAMTSSINT"  
 BASE COUNT 196 a 351 c 299 g 255 t  
 ORIGIN  
 Query Match 11.5%; Score 234.8; DB 4; Length 1101;  
 Best Local Similarity 69.3%; Pred. No. 9.5e-27;  
 Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
 QY 92 GCGCCTTTCCCTGGGGGCGCTGCGCGGAGCGCTGTGCGCTGCTGCTGCTGCTG 151  
 Db 107 CGCTTTCCTCCAGCGCGCTGTGGCGGCGTACCGACCTGCGCTGCTGCTGCTG 166  
 QY 152 TCGGGGTAGGCGCAAGCTGTGAGCTGATGCTGATCGGGGCGTACCGGAGATCGGA 211  
 Db 167 TGGGTATGCGGGCAACCTGCTCAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 226  
 QY 212 CCACCAACAATTGTAACCTGCGGAGATGCGCTGTGCTGCACTTCACTTCTGCGGGC 271  
 Db 227 CCACCAACAACCTTACCTGCTGCGAGATGCGCTTCTCCAGCACTTCACTTCTGCA 286  
 QY 272 TGGCGTTGACCTGTACCGGCTTCTGCGGCTGCGGCGCTGCGGCTGCGGCTGCTCT 331  
 Db 287 TGGCGCTGACCTTCTTCCGCTTCTGCGAGTACCGGCTTGTGAACCTTGTGCAACTGCTCT 346  
 QY 332 GCGGCTGCTGCTTACGCTGAGGCGAGGCTGCACTTACCGCAAGCTGTGACATGACCG 391  
 Db 347 GCAACTCTTCCAGTTCGTGTAGAGAGTGTGCACTTACCGCAAGCTGTGACATGACCG 406  
 QY 392 CGCTGAGCGTGCAGCGCTACCTGCGCATCTGCGCGCTGCGCGCTGCGCGCTGCTGTG 451  
 Db 407 CGCTGAGCGTGCAGCGCTACCTTCCGCTGCGCATCTGCTTCCGCTGCGCGCAAGTGTG 466  
 QY 452 CCGGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCG 511  
 Db 467 CCAAGGCGCGGTAAGCTGTGATCTGTGTATCTGCGCGCGCTGCTGCTGCGCGCG 526  
 QY 512 GTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553  
 Db 527 GAGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568

Search completed: March 16, 2003, 03:14:23  
 Job time : 6757 secs